

Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech

Jennifer Cole^{a,*}, Heejin Kim^a, Hansook Choi^a, Mark Hasegawa-Johnson^b

^aDepartment of Linguistics, 4088 Foreign Languages Building, 707 South Mathews, University of Illinois at Urbana-Champaign, Urbana, IL 61822, USA

^bDepartment of Electrical & Computer Engineering, 155 Everitt Lab, 1406 W. Green, University of Illinois at Urbana-Champaign, Urbana, IL 61822, USA

Received 1 April 2004; received in revised form 25 March 2006; accepted 27 March 2006

Abstract

This study examines the effects of prosody on the acoustic cues of stop consonants, based on speech from the Boston University Radio News corpus. An investigation of VOT, f_0 , closure duration, burst amplitude and spectral characteristics provide evidence for a primary effect of accent (a level of phrasal prominence) on these measures as cues to stop voicing and/or place of articulation. There are robust, significant effects of accent on the acoustic cues for both voicing and place of articulation, exhibiting a general pattern of prosodic strengthening. However, the accent effect on voicing cues results in the enhancement of the phonological voice contrast, whereas the effect on acoustic cues to place of articulation does not similarly enhance the place contrast. Comparison of η^2 values shows that accent effects are strongest on those measures that are weak cues to the phonological contrast, and weakest on the measures that are strong cues. While there are no observed strengthening effects of phrasal position on voicing cues for /t/ and /d/, acoustic variability is reduced in phrase-initial position.

© 2006 Elsevier Ltd. All rights reserved.

1. Introduction

Through the modulation of f_0 , intensity, and duration, a speaker conveys information about the grouping of words into phrases, the information status of words and phrases, and about the speaker's intentions and emotions related to the utterance and the discourse situation. Though some of these acoustic dimensions, notably f_0 and duration, may also be used to signal lexical contrasts in a language, their role in marking the structural and pragmatic content of an utterance is distinct and defines the prosody of a language. An interesting finding from recent work is that the suprasegmental features encoded in f_0 , intensity and duration are not the only manifestations of prosody. Prosodic phrasing and accentuation also condition the way in which the speech sounds in the word are implemented in articulation, thus giving rise to prosodically conditioned acoustic variation.

*Corresponding author.

E-mail address: jscole@uiuc.edu (J. Cole).

Evidence for the effects of prosody on phonetic variation is found in numerous studies based on “laboratory” speech, typically consisting of short scripted utterances produced in the absence of a meaningful communicative context. Rather less is known about the extent to which prosody influences phonetic variation in speech produced in meaningful contexts. On the one hand, we might expect prosodic cues to phrasing and prominence to be greater in non-laboratory speech. In real discourse contexts the speaker presumably has a greater concern that his or her message be correctly perceived by the listener, and may be expected to use the full range of prosodic expression to signal the phrasing and pragmatic content of the utterance. On the other hand, in a natural discourse context the identification of the words in the utterance, along with their phrasing and pragmatic content is to some extent guided by sentential and discourse context. The speaker may not need to exploit the full range of prosodic expression to be correctly understood, and considerations of speech economy may dictate prosodic reduction, along the lines suggested by Lindblom (1990) for segmental reduction.

In this paper we report on evidence for the effects of prosody on radio news announcers’ speech (American English), based on an acoustic analysis of the speech of four professional news announcers from the Boston University Radio News corpus (Ostendorf, Price, & Shattuck-Hufnagel, 1995). Our study is focused on oral stop consonants and the acoustic properties that mark the phonological contrasts of voicing and place of articulation (POA). Through measurements of duration, energy and the spectral envelope, we look at the effects of prosody on the acoustic correlates of stop voicing and POA. The primary prosodic factor we consider in the present study is accent, where this term is used to indicate a level of sentential stress prominence that is higher than lexical stress and which is accompanied by a salient pitch accent (Beckman & Pierrehumbert, 1986). We also consider in lesser detail (due to data limitations) the effect of prosodic position by comparing the stops/t,d/ in the initial vs. medial positions in the intonational phrase (IP). Section 2 reviews evidence from prior research for prosodic effects on articulatory and acoustic variation and presents the hypotheses for our experiments. Section 3 presents our acoustic study of Radio News speech, establishing the acoustic correlates of voicing and place of articulation in these data, and demonstrating the significant strengthening effect of accent on some of these acoustic measures. Our preliminary results on the effect of phrase-initial position fail to show a similar pattern of positional acoustic strengthening. Section 4 discusses the pattern of results, drawing comparisons of effect size between the phonological factors of Voice and POA and the prosodic factors of Accent and Position. The results presented here are shown to provide only partial evidence for a pattern of phonological contrast enhancement through prosodic strengthening. For some acoustic measures, the effect of prosodic strengthening results in a greater distinction between phonologically contrastive sounds, while other measures show strengthening effects that do not serve to enhance contrast.

2. Prosodically conditioned phonetic variation

Instrumental studies of speech articulation show two primary effects of prosody in varieties of English. First, speech gestures are strengthened in syllables that are lexically stressed or marked with phrasal accent (Beckman & Edwards, 1994; Beckman, Edwards, & Fletcher, 1992; Cho, 2004, 2005, 2006; de Jong, 1991, 1995; Edwards, Beckman, & Fletcher, 1991; Erickson, 2002; Harrington, Fletcher, & Beckman, 2000; Harrington, Fletcher, & Roberts, 1995), or in initial position in the phrasal domain (Cho, 2004, 2005, 2006; Cho & Jun, 2000; Fougeron & Keating, 1997). Strengthening may affect speech gestures in various ways, including greater spatial displacement, greater velocity and/or longer duration. Second, speech gestures are lengthened in domain-final position (Beckman et al., 1992; Byrd, 2000; Cho, 2006; Edwards et al., 1991). Final lengthening affects the duration of the speech gesture, with conflicting evidence about the presence (Cho, 2001, 2005, 2006; Tabain, 2003) or absence (Beckman et al., 1992; Edwards et al., 1991) of simultaneous gestural strengthening.

The articulatory effects of prosody described above suggest the presence of corresponding effects on acoustic properties of speech sounds. To a limited extent, acoustic evidence confirms this expectation for English. Acoustic effects of accent are reported in the increased duration of vowels (Edwards et al., 1991) and consonants (Turk & White, 1999), similar to the lengthening effects reported for lexical stress (Crystal & House, 1988b; Klatt, 1974; Umeda, 1977). Pierrehumbert and Talkin (1992) report acoustic evidence for the accentual strengthening of /t/ in terms of increased voice onset time (VOT), and for /h/ in reduced energy and

in spectral measures of the ratio of the first and second harmonics. Choi (2003) reports increased VOT and f_0 values under accent for the stops /p,b/, with the greatest effect on voiceless /p/. There is also evidence of the effect of accent on spectral measures, as reported by Wouters and Macon (2002), with accent-induced increases in the spectral rate of change in the transitions between vowels and adjacent liquids and glides. Regarding acoustic effects of prosody on vowel quality, Erickson (2002) reports effects of emphasis (corresponding to pitch-accent) on F1 and F2 measures, with an increased distinction between non-low, front vowels and low vowels under accent. Cho (2005) shows evidence of vowel strengthening under both accent and boundary conditions, based on F1 and F2 measures from the vowels /a, i/. Further acoustic evidence of strengthening in initial position of a prosodic domain is found in increased consonant duration (Fougeron & Keating, 1997), and in the increased likelihood of glottalization for vowels that are initial in higher prosodic domains compared with vowels that are initial in domains lower in the prosodic hierarchy (Dilley, Shattuck-Hufnagel, & Ostendorf, 1996). Keating (1984) finds evidence of increased VOT for voiced and voiceless stops in domain-initial position (words in citation form), though we note that Choi (2003) fails to find consistent evidence of strengthening in domain-initial position (IP) on VOT or f_0 values for /p,b/. Acoustic evidence of domain-final lengthening is found in Wightman, Shattuck-Hufnagel, Ostendorf, and Price (1992) for all segments in the domain-final syllable rime, and in Edwards et al. (1991) for the vowel of the domain-final syllable.

In addition to the studies mentioned above on prosodic effects on English, similar studies investigate prosodic effects in other languages, e.g., Korean (Cho & Jun, 2000; Cho & Keating, 2001; Jun, 1993), French (Fougeron, 2001; Keating, Cho, Fougeron, & Hsu, 2004; Tabain, 2003), Tamil (Byrd, Kaun, & Narayanan, 1996), Dutch (Cho & McQueen, 2005), and Taiwanese (Hsu & Jun, 1998; Keating et al., 2004). Although the details differ somewhat between languages, the findings are generally consistent with the prosodic strengthening and lengthening effects observed in English.

With one exception, the studies cited above are based on data from “lab speech”: highly controlled speech that is selected by the experimenter for the purpose of the investigation, and read by the speaker in a laboratory setting. The exception is the Dilley et al. (1996) study, which is based on the same corpus of radio news speech that is used in the present study. We also note here the study by Greenberg, Chang, and Hitchcock (2001) that finds a lengthening effect of sentential stress on vowel duration in the Switchboard corpus of conversational telephone speech. Related work on English by Umeda (1977) and Crystal and House (1988b) examines the effect of lexical stress on the duration of speech sounds based on short readings of an essay or story prepared by the experimenter. This material provides more of a communicative context than is typical for lab speech, but differs from Radio News speech in the experience level of the speaker(s) in reading scripted speech, and in the presence of a genuine communication task. Note that neither of those studies look at prosodic features involving phrase-level stress or accent, or prosodic domains. In addition to these works on English, there is also the study by Duez (1992) on French, notable for its use of spontaneous speech data, that looks at the effect of sentential prominence on locus equation measures of CV coarticulation.

The strengthening effect conditioned by prosody sometimes results in an enhancement of the phonological contrast between two or more sounds that belong to the same major sound class (C or V); this is conventionally termed paradigmatic strengthening.¹ Evidence of prosodically conditioned contrast enhancement has been reported from several studies. Cho (2001, 2004) argues that articulatory strengthening in accented and domain-boundary positions increases the phonetic clarity of speech sounds, thereby enhancing their distinctiveness. In their acoustic study of VOT as a cue to stop aspiration and voicing in Taiwanese, Hsu and Jun (1998) find evidence of contrast enhancement in greater positive VOT values for aspirated /k^h/ and greater negative VOT (voicing lead) for /b/ in domain-initial position, with the greatest distinction in initial position of the IP. Choi (2003) finds similar evidence of contrast enhancement in the acoustic cues to stop voicing in American English under focal accent.

¹Here, as elsewhere in the literature, the term “contrast enhancement” is used to describe the increased distinctiveness of two or more sounds in acoustic space, and does not explicitly refer to any measurement of the perceptual distinctiveness of the sounds in question. While we do not assume a simple linear relationship between acoustic and perceptual distinctiveness, we conjecture that increased acoustic distinctiveness will be positively correlated with increased perceptual distinctiveness for any acoustic property that has been shown to play a role in the perception of a phonological contrast.

On the other hand, it is possible to have prosodic strengthening without phonological contrast enhancement. For instance, if prosodic strengthening affects both members of a phonologically contrastive pair in the same way, i.e., with a uniform increase or decrease of a relevant acoustic measure, then the prosodic strengthening will not improve the distinction between the contrastive sounds, and therefore will not generally serve to enhance the phonological contrast. We will use the term **uniform strengthening** here to refer to such a pattern. Under uniform strengthening, an acoustic property that marks a phonological contrast may be rendered more salient overall (e.g., through longer duration or higher amplitude), but in a manner that does not increase the distinctiveness of the phonologically contrastive sounds with respect to that property. We distinguish uniform strengthening from the pattern of prosodic strengthening that results in **contrast enhancement**, as illustrated in Fig. 1, which provides a schematic illustration of four possible effects of prosodic strengthening on mean values of some acoustic measure for sounds belonging to two hypothetical phonological categories, for example, voiced stop vs. voiceless stop. The chart in (a) illustrates contrast enhancement, with an increase in the acoustic measure of category 1 sounds under Accent, and a decrease in the same measure for category 2 sounds under Accent. The contrast between categories 1 and 2 is enhanced in accented positions relative to unaccented positions because the distance between the sounds in this acoustic dimension is increased. The chart in (b) illustrates uniform strengthening, because sounds from both categories are affected by Accent in the same way showing similar patterns of increase in the acoustic measure, and are not more distinct from one another (i.e., not more distant in this acoustic dimension) in accented positions relative to unaccented positions. Chart (c) presents an interesting scenario, because sounds in both categories are affected by Accent in a similar way; both undergo increases in the relevant acoustic measure, but because the increase is much larger for the category 1 sounds, the result is contrast enhancement under Accent. Thus, both pattern (a), with opposite acoustic effects for categories 1 and 2, and pattern (c), with acoustic effects in the same direction for categories 1 and 2, provide evidence of contrast enhancement. Chart (d) presents another case of asymmetric strengthening without contrast enhancement. This is similar to pattern (c), but in (d) category 2 increases more than category 1 under Accent, with the result that contrast is actually diminished under Accent. The sections below present evidence for all of these patterns, revealing uniform strengthening, contrast enhancement and even a few instances of diminished contrast under Accent.

Another pattern of prosodic strengthening effect has been noted in the literature, where strengthening serves to enhance the distinction between consonants and vowels in a way that presumably facilitates segmentation

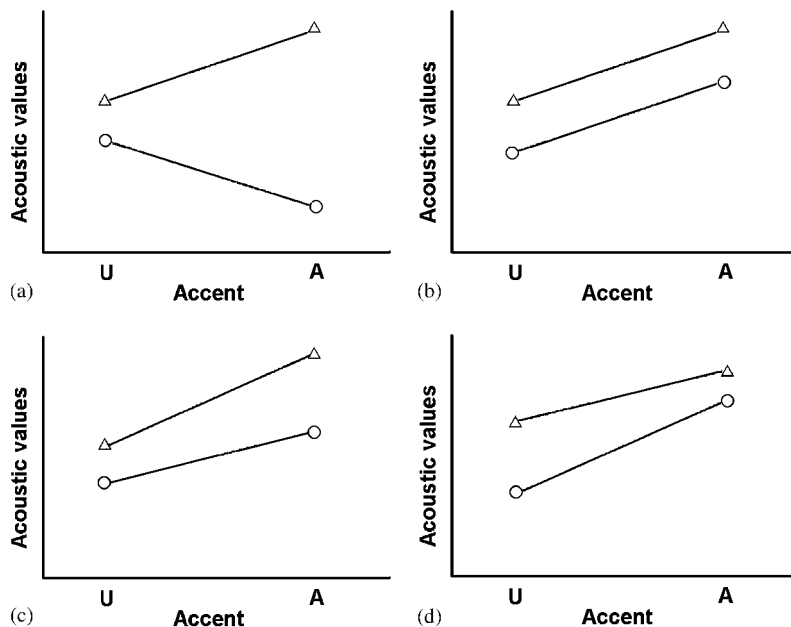


Fig. 1. Schematic graphs illustrating the effect of prosodic strengthening under Accent on some acoustic measure, for sounds that belong to two phonologically contrastive categories: Δ , Category 1; \circ , Category 2, in Unaccented (U) and Accented (A) positions.

of the speech signal into a sequence of discrete sounds; this is traditionally termed syntagmatic strengthening. Under syntagmatic strengthening consonants exhibit increased constrictions while vowels have increased opening gestures (Beckman & Cohen, 2000; Beckman et al., 1992; de Jong, 1995; Fougeron & Keating, 1997; Jun, 1995 (for Korean); Pierrehumbert & Talkin, 1992). In this manner consonants and vowels display increased distinctiveness relative to one another.

In our study of stop consonants in Radio News speech, we examine the acoustic properties that encode stop voicing and place of articulation to look for evidence of prosodic strengthening under accent or phrase-initial position. The speech of professional radio announcers is described as being clearer with more consistent indications of prosodic structure than non-professional read speech (Ostendorf et al., 1995), and thus we expect to find clear indications of prosodic effects at the segmental level in this type of speech, as have been described for laboratory speech. Further, given the findings from the articulatory studies cited above that speech gestures are faster, longer and of greater magnitude in prosodically strong positions (accent, phrase-initial), we expect to find some acoustic evidence for this pattern of strengthening in Radio News speech. Specifically, we focus on the voicing and place contrasts among the set of stop consonants and look for evidence of strengthening that results in contrast enhancement. A secondary focus is on evidence for syntagmatic strengthening in the acoustic cues to stop closure and release that would serve to discriminate a stop from an adjacent vowel. The following hypotheses guide our experimental work:

Hypothesis 1. Prosodic strengthening enhances the Voice contrast. The distance between voiced and voiceless stops in acoustic space will be greater in prosodically strong positions (accented, phrase-initial) than in weak positions (unaccented, phrase-medial).

The acoustic cues to phonological voicing measured in this study are VOT, f_0 at onset of the following vowel, Closure Duration (CD), and Burst Amplitude (BA). In American English, VOT values are higher for voiceless stops and lower or near zero for voiced stops. f_0 values at the onset of the following vowel (or sonorant consonant) are lower for voiced stops and higher for voiceless stops (Lisker & Abramson, 1964; Whalen, Abramson, Lisker, & Moody, 1993). The evidence for a voicing effect on stop CD is inconsistent. Chen (1970) reports longer CD for voiceless stops in citation forms, but Luce and Charles-Luce (1985) find only sporadic evidence of a voicing effect on CD in phrase-final position, and conclude that CD fails to distinguish voicing consistently across positions and local phonetic context. Crystal and House (1988a) similarly fail to find clear evidence of a voicing effect on CD. Lisker (1957) and Port (1976, 1979) find evidence of a CD distinction for voiced and voiceless stops, but observations are restricted to stops in intervocalic, post-stress position, as in the pair *rabid*, *rapid*, where the voiceless stop has longer CD. Among these prior works, the speech materials from which our data are collected most closely resemble the read speech of the Crystal and House (1988a) study, and so our expectation for a voicing distinction on CD measures is correspondingly diminished. Finally, with respect to BA, voiced stops are typically found to have lower BA than voiceless stops (Zue, 1976). BA differences between voiced and voiceless stops may be aerodynamically related to the length of the oral closure, along the line of reasoning put forth by Cho and Keating (2001) to account for VOT differences related to CD. Oral pressure increases with increased CD, and if the rate of release is the same for voiced and voiceless, then we may expect to observe higher BA values on the basis of longer CD with voiceless stops.

In addition to the studies cited above documenting the acoustic correlates of stop voicing, there are many studies that provide evidence for the role of these acoustic properties as cues for the perception of stop voicing. Liberman et al. (1958) and Lisker and Abramson (1967) establish the role of VOT as a cue to stop voicing in perception studies using synthesized speech stimuli. Later work by Stevens and Klatt (1974), Summerfield and Haggard (1977) and Whalen et al. (1993), also using synthesized speech stimuli, confirm the role of VOT as a voicing cue, and demonstrate an interaction between VOT and f_0 or F1 onset frequency. The Whalen et al. (1993) study also contributes evidence for the role of f_0 as a voicing cue, even when VOT is unambiguous, confirming earlier findings on f_0 by Haggard, Ambler, and Callow (1970). Repp (1979) presents evidence for BA as a perceptual cue to voicing, although the effects of BA on voicing perception are found to be lesser than the effects of aspiration amplitude. Finally, Port (1976, 1979) finds evidence for CD as a perceptual cue to stop voicing.

With each of the four voicing cues mentioned above, VOT, f_0 , CD and BA, voiced and voiceless stops are distinguished by the range of values they take: high or low relative to the opposing voice category. Prosodic strengthening is expected to enhance the contrast between the voiced and voiceless categories with respect to one or more of these cues, by increasing the values of the high category, or by decreasing the values of the low category, or through both effects in combination. Thus, in prosodically strong positions, under the hypothesis of prosodic strengthening we expect to find: voiceless stops with increased VOT, f_0 , CD or BA; or voiced stops with decreased VOT, f_0 , CD or BA; or some combination of these effects.

Hypothesis 2. Prosodic strengthening enhances the Place contrast. The distance between labial, alveolar and velar stops in acoustic space will be greater in prosodically strong positions than in weak positions.

In the present study we examine multiple cues to stop place of articulation. The VOT, CD and BA measures considered above as cues to the voicing contrast have also been found to distinguish place of articulation, as follows.

VOT: Labial < Alveolar < Velar (Crystal & House, 1988a; Klatt, 1975; Plauché & Sönmez, 2000).²

BA: Labial < Velar < Alveolar, where the difference between the Labial and Velar groups occurs depending on the following vowel context (Ohde & Stevens, 1983; Plauché & Sönmez, 2000).

CD: Velar < Alveolar < Labial (Luce & Charles-Luce, 1985; Umeda, 1977), though Crystal and House (1988a) report Alveolar < Labial, Velar.

Evidence for the role of these acoustic measures as cues to the perception of stop POA is reported in Massaro and Oden (1980) and Benki (2001) for VOT; Ohde and Stevens (1983) for BA; and Tartter, Kat, Samuel, and Repp (1983), Dorman and Raphael (1980) and Port (1976) for CD.

In prosodically strong positions, the differences between labials, alveolars and velars in these acoustic measures are expected to be bigger, through increases in the high-valued category, decreases in the low-valued category, or a combination of both. The middle-valued category may also vary with prosodic factors, but the direction of the variation is not predicted by the hypothesis of prosodic strengthening.

In addition to the energy and duration cues just mentioned, POA is also known to be encoded by the spectral envelope of the stop release burst (Blumstein & Stevens, 1979). Labials are described as having a burst spectrum with greater peak amplitude in the lower frequency region (diffuse-falling), alveolars as having greater peak amplitude in the high frequency range (diffuse-rising), and velars as having a compact spectrum with characteristic peak amplitude in the middle frequency region. Evidence for the burst spectrum as a perceptual cue to stop POA is reported in Winitz, Scheib, and Reeds (1972), Stevens and Blumstein (1978), and Repp and Lin (1989). While the work by Steven and Blumstein looks at static properties of the burst release, spectral dynamics are also found to be significant cues to stop place in the subsequent work by Kewley-Port (1983) and Lahiri, Gewirth, and Blumstein (1984). The present study looks only at static measures of the release burst in the Radio News speech data, leaving open the possibility that further prosodic effects may exist in the spectral dynamics of the burst release.

The location of the spectral peaks is determined jointly by the dimensions of the vocal tract and the location of the constriction in the vocal tract. Prosody could affect the former, for example, through strengthening of labial gestures that effectively lengthen the vocal tract under accent, though effects of this sort have not been attributed to prosodic factors in prior work. Prosodic strengthening could also affect constriction location, in that a hyper-articulated accented stop may be more likely to reach its target constriction location compared to the unaccented stops that are produced with reduced gestures (hypo-articulated). In this scenario, there should be less variation in the spectral measures of POA for stops in prosodically strong positions compared to stops in weaker positions.

²In addition to these studies on VOT in American English, there are several acoustic studies on other language varieties showing that VOT values of stops vary according to POA, notably Docherty (1992), Jessen (1998), and Cho and Ladefoged (1999). Docherty (1992), in his study of VOT in British English, finds a distinction between labials and non-labials, but does not find a robust VOT difference between alveolars and velars. A similar pattern is found in Cho and Ladefoged's study of 18 languages (not including English), where significant differences are found in mean VOT values between velars and coronals (dentals and alveolars), but bilabials and coronals are not significantly different. It is notable that only the studies on American English report a reliable difference in VOT between labials and alveolars.

Hypothesis 3. Prosodic strengthening enhances the C/V segmental boundary (**syntagmatic strengthening**). The juncture between a consonant and an adjacent vowel as measured in the acoustic correlates of stop closure and release will be enhanced (uniformly for all stops) in prosodically strong positions compared to weak positions.

This enhancement will be evident in the acoustic cues that mark consonant duration and the constriction release. Specifically, both voiced and voiceless stops in strong positions are expected to exhibit greater positive VOT and longer CD to increase the duration of the consonantal region with lower or non-periodic energy, and higher BA to more distinctly mark the juncture between the consonant and the onset of voicing for the following vowel. We do not expect any significant effects of syntagmatic strengthening on the spectral place cues; it is not evident that changes in the location or variability of spectral peaks would enhance the distinctiveness of the consonant constriction in the context of an adjacent vowel. A consonant that is produced with greater compression under prosodic strengthening may exhibit small shifts in the location of spectral peaks as the edges of the constriction advance to cover a larger area, but this shift would not necessarily increase the distinctiveness of the consonant relative to an adjacent vowel.

Another effect of syntagmatic strengthening might obtain for consonants in word-initial position, where strengthening could serve to enhance the salience of the word boundary. The current study looks only at data from stops in word-initial position, and so can not investigate the effect of strengthening at word boundary compared to word-internal stops.

3. Method

3.1. Corpus

The speech analyzed in this study comes from the prosodically labeled “lab news” portion of Boston University Radio News corpus (Ostendorf et al., 1995). Four news stories that had been earlier broadcast were transcribed and read by professional radio news announcers in a laboratory setting, with the announcers speaking in their radio style. This portion of the corpus includes six speakers in total, out of which we analyze four here: F3A, F2B, M1B, M2B. The speaker labels “F” and “M” designate female and male speakers, the label “A” designates that the announcer normally reads news live, and “B” designates that the announcer normally pre-records and edits their stories. The speakers are reading the same four news stories, which contain a combined total of 119 sentences and 2123 words. The amount of lab news data that is prosodically labeled differs for each speaker, but consists of a total of about 300 sentences for the four speakers combined, for about 42 min of speech, sampled at 16 kHz. Note that some of the sentences in this corpus are long and include conjunctions and/or disjunctions, and may comprise several complete IPs.

The lab news speech is accompanied by word transcripts using the TIMIT set of phone labels, and a prosodic transcription based on the ToBI labeling standard (Beckman & Elam, 1997) that marks the location and tonal melody of pitch accents and phrasal boundaries, based on the labeler’s integration of auditory cues with visual cues from the waveform, spectrogram and F0 track. Labeling criteria are described in detail in Beckman and Elam (1997). As reported by Ostendorf et al. (1995), the phone labels and segmentation were generated automatically, and the prosodic transcription was manually produced by trained labelers. Pitch accents are transcribed on the stressed syllable of a word that is perceived as prominent in the phrase and that realizes a marked F0 excursion. Seven types of pitch accents are distinguished by their F0 contour: H*, !H*, L+H*, L+!H*, L*, L*+H and H+!H*. L and H correspond to tonal targets, with an asterisk marking the tone within a complex (falling or rising) melody that is aligned with the stressed syllable, and !H indicates a downstepped high tone. Phrasal boundaries are marked with juncture labels that appear at the end of a word that is final in the phrase. Two levels of phrasal juncture are marked in the transcription: level 3 marks an intermediate phrase boundary and level 4 a full IP boundary. The juncture labels 0–2 are transcribed to mark lower levels of juncture between words that are phrase-medial, but are not considered here. The phrasal junctures are further marked with tones based on the F0 contour at the phrase edge: L-, H- or !H- for intermediate phrase junctures, and L% or H% for IP junctures. Phone and prosody labels were not subject to re-evaluation in the present study, but incorrect phone and prosody alignments were manually corrected whenever noticed.

The stop consonants from the set /p t k b d g/ were located in the dataset based on word transcriptions and were manually segmented on the basis of waveform, spectrogram and listening, using standard conventions for the location of acoustic landmarks marking the left and right edges of the stop (Crystal & House, 1982). Each stop was coded for its lexical stress on the basis of the dictionary form, and for its accentuation (accented, unaccented) as determined by the presence or absence, respectively, of a pitch accent label on the syllable that contains the stop. Stops were also coded for prosodic position relative to the IP, the highest level juncture specified in the prosodic transcription. Phrasal encoding distinguished stops in phrase-initial position from those in phrase-medial position. Phrase-final stops were also identified, but the uneven distribution of the various stops in final position prevented comparison with initial and medial stops, and so those data are not reported here. All stop tokens are word-initial and pre-vocalic (#CV). Stops that occur in word-initial consonant clusters are not included in this study.

Since the stop tokens were collected from a speech database that was not controlled for the purpose of this experiment, the number of tokens of each stop under each prosodic condition is not equal. In particular, there were many fewer stop tokens in IP-initial position than IP-medial, which is not surprising given that an IP typically contains 4–10 words out of which only one word is in initial position. Similarly, the data were not controlled for the quality of adjacent vowels, so effects due to specific vowel contexts are not investigated here. Generally speaking, the frequency of vowels adjacent to the stop consonants studied here mirrors the overall frequency of vowels in this corpus. As an approximation of the distribution of vowels in the contexts of stop consonants, Table A1 in the Appendix gives the frequency of individual vowels (based on lexical values) following the studied stop consonants for speaker F3A, along with counts of the number of stop tokens from function words. Note that there is a high occurrence of /b/ and especially /t/ tokens due to their occurrence in high frequency function words like “by” and “to”; all but four of these function word tokens are in unaccented words. The decision to include data from function word contexts was based on a preliminary analysis of data from speaker F3A that compared the results for acoustic measures between the larger data set that includes function words with a smaller data set from which function words were removed. There were no substantive differences between these two data sets in the means and distributions of acoustic measures in accented or unaccented contexts. Although significance levels and effect size values varied by small degrees,

Table 1

Distribution of stop tokens for each speaker according to prosodic condition: IP-initial unaccented (IU); IP-initial accented (IA); IP-medial unaccented (MU); IP-medial accented (MA). All stop tokens occur word-initially and are followed by a vowel

Speaker	Prosodic condition	p	b	t	D	k	g
F3A	IU	1	12	13	9	3	0
	IA	1	1	1	02	5	1
	MU	19	40	59	21	32	14
	MA	17	24	17	9	24	5
	Total	38	77	90	41	64	20
F2B	IU	1	19	17	12	5	3
	IA	4	1	4	4	6	0
	MU	23	39	44	21	37	13
	MA	7	21	17	6	17	3
	Total	35	80	82	43	65	19
M1B	IU	1	14	8	5	2	0
	IA	1	0	2	4	4	0
	MU	14	51	67	29	31	14
	MA	18	21	18	10	28	7
	Total	34	86	95	48	65	21
M2B	IU	1	12	10	9	4	0
	IA	3	3	4	3	5	1
	MU	17	32	34	22	26	10
	MA	16	26	17	14	21	8
	Total	37	73	65	48	56	19

there were no differences in main effects, and relative effect sizes were also the same. The function word data, which are nearly all from unaccented contexts, very closely resemble the unaccented, non-function word data. Based on this finding, the results reported here include data from function word contexts.

The distribution of stops in each prosodic condition for each speaker is given in Table 1. The number of tokens that occur in IP-initial position is too few to allow a complete analysis of the acoustic cues to either voicing or place of articulation in initial position compared to medial position. For this reason we focus mainly on the prosodic effect of accent in this paper, and limit our remarks on the effect of position to observations of /t/ and /d/, for which adequate data is available.

3.2. Measurements

The stop tokens were manually labeled at acoustic landmarks for the subsequent extraction of the measurements listed below. Acoustic analysis was performed using Praat, a software application for speech analysis (Boersma & Weenink, 2006).

VOT: VOT was measured from the first peak of the stop release burst up to the zero crossing nearest to the onset of the second formant of the following vowel. Some tokens of the voiced stops /b d g/ were produced with closure voicing, and for those tokens where closure voicing extended up to the stop release, no VOT interval was measured and the token was labeled as “prevoiced”. The number of prevoiced stops for each speaker is presented in Section 3.3.1 along with the VOT results.

F₀: F₀ measures at the onset of voicing following stop release where estimated from the mean duration of the first three periods, based on visual inspection of the waveform, and confirmed by observation of the vertical striations corresponding to glottal pulses in the spectrogram. This measure was compared to the f₀ value obtained using the autocorrelation method of Praat with a time step of 10 ms., and although the two measures were found to be very close, the automatic measurement sometimes failed to assign an f₀ value at voicing onset for tokens that could be successfully analyzed manually, and so the manual measurement was adopted for all tokens.

With regard to CD, for stops following a liquid or vowel (in the preceding word), the beginning of the closure interval was marked at the point where the second and higher formants end. When the preceding sound was a fricative, the left side of the CD was marked at the end of frication noise. When the preceding sound was a stop, the left side of CD was marked after the stop release, and if there was no stop release then no CD measurement was taken. The right edge of the closure interval was marked at the beginning of the stop release. Stop tokens that occur in IP-initial position are frequently preceded by a silent pause and hence were not labeled for CD, as the beginning of the closure interval was not separable from the preceding silence.

For BA, measures were based on RMS intensity measured in a 10 ms. Gaussian window was aligned at the left edge with the first peak of the burst, as visually identified on the waveform display.

Burst Spectrum: Measurements of burst spectrum were made to assess the distribution of energy in the release burst as a function of the stop place of articulation. RMS intensity was measured in three frequency bands resulting from passband filtering at frequencies of 0–1 kHz, 1–3 kHz, and 3–5 kHz, in the same 10 ms window used for the BA measurement. These measurements are referred to below as BS–low, BS–mid, and BS–hi. Along with the intensity measures for each frequency band, we also looked at two ratio measures of intensity: mid/low (1–3 kHz/0–1 kHz) and mid/high (1–3 kHz/3–5 kHz). These measurements are referred to below as BSRatio–low and BSRatio–hi, respectively. These ratios are expected to distinguish place of articulation as follows. The diffuse-falling spectrum of the labials should be distinguished by mid/low ratio values less than or equal to 1, and mid/high ratio values greater than or equal to 1. The diffuse rising spectrum of alveolars should yield the reverse, with mid/low ratios near or greater than 1, and mid/high ratios less than 1. The compact spectrum of velars with its characteristic mid frequency peak should yield comparable mid/low and mid/high ratios, both much greater than 1.

3.3. Results

The results are presented in five sections, beginning with the correlates associated with Voice and POA distinctions on IP-medial stops in Section 3.3.1. The IP-initial stops are excluded from this analysis due to the

limited occurrence of some stops phrase-initially, making broad comparisons across Voice, POA or Accent groups unreliable. Section 3.3.2 presents results for the effect of Accent on VOT, f_0 , CD and BA as cues to Voice and POA for IP-medial stops, and Section 3.3.3 discusses the interaction of lexical stress and phrasal accent in the results for Voice cues for one speaker. Section 3.3.4 presents results for the effect of Accent on burst spectrum cues to POA, and Section 3.3.5 presents a limited comparison of initial and medial /t/ and /d/ for the effects of IP-position on the acoustic cues to voicing for this pair, the only stops for which there were sufficient observations in initial position to license statistical analysis.

In Section 3.3.1 the correlates of Voice and POA distinctions on phrase-medial stops are considered for the acoustic measures of VOT, f_0 , CD, BA and burst spectral energy to establish the pattern of cues used by each speaker to encode the phonological Voice and POA contrasts. The dependent acoustic measures are subject to 2-way ANOVA for each speaker individually, with the independent factors Voice (Voiced, Voiceless) and POA (Labial, Alveolar, Velar). Alongside the ANOVA results are the Partial η^2 values for the estimates of effect size, which describe the portion of the variance accounted for by the independent factor. η^2 values range from 0 to 1, with larger values indicating stronger effects. ANOVA is generally robust to differences in variance across groups, as long as the groups are roughly equal in size. In this study many of the groups being compared are of unequal size. Levene's test for homogeneity of variance is applied to identify cases where comparison groups have unequal variance. Those cases are subject to the non-parametric Kruskal–Wallis test for differences between groups and results of the Kruskal–Wallis test are also reported for cases where the ANOVA and non-parametric results differ. Post hoc analysis is performed with the Tukey and Tamhane's T2 tests for the 3-level factor of POA. The Tamhane's T2 test makes no assumption of equal variance across groups, and is used here based on the findings of Levene's test.

The results for the effect of Accent on the dependent acoustic measures for medial stops, presented in Section 3.3.2, are subject to 2-way ANOVA with the independent factors Accent and Voice first, and then Accent and POA. Section 3.3.5 similarly presents the results for the effect of IP-position on the acoustic voicing cues for /t/ and /d/, based on 2-way ANOVA with the factors Position and Voice. The Position effect is analyzed first for only unaccented /t,d/, and then for the combined accented, unaccented /t,d/ data. An analysis of the interaction of Accent and Position effects for /t,d/ is not possible due to data limitations.

3.3.1. Acoustic correlates of Voice and POA distinctions in IP-medial stops

Descriptive statistics are provided in Table A2 of the Appendix showing the mean and standard deviation for each of four acoustic cues to Voicing—VOT, f_0 , CD, BA—for each speaker individually. The number of missing values (resulting from inability to identify and extract the relevant acoustic measures from the signal) is also given for each acoustic measure, as an index of how robust the measure is for the given category and speaker in this speech corpus. Some of these acoustic measures also encode POA contrasts for some speakers, and Table A3a,b of the Appendix provides descriptive statistics for the same four acoustic measures, this time grouped by POA. For all speakers there are many missing items from the two durational measures, VOT and CD. Missing VOT values usually occur with voiced stops when the speaker produces closure voicing that extends up until the release (prevoicing). In such cases there is no voiceless region to mark the left edge of the VOT. Data on the duration of closure voicing is not reported here, because for a large number of these phrase-medial tokens the beginning of the stop voicing interval could not be reliably segmented from the preceding voiced segment, a problem which is also reflected in missing CD values for those tokens. There is a lot of inter-speaker variation in the production of voiced stops with closure voicing, with speaker F3A producing the greatest number. The missing CD values come mostly from tokens where the target stop is preceded by an unreleased stop, in which case the beginning of the closure interval cannot be reliably segmented from preceding context.

The results of 2-way ANOVA with the independent factors of Voice and POA are shown for each speaker in Tables 2a,b. A significant effect of Voice is seen on VOT and f_0 for each speaker. Two speakers (F3A, M2B) show a significant effect of Voice on CD and two speakers (F2B, M2B) show a significant effect on BA. High η^2 values for the Voice effect on VOT (F3A: 14.8%; F2B: 47.6%; M1B: 45.7%; M2B: 26.1%) indicate that Voice accounts for much of the variation in VOT. A comparison of η^2 values across the four acoustic measures shows that Voice accounts for a much greater portion of the VOT variation than it does for the variation of other acoustic measures, for all speakers. The Voice effect on BA for speaker M2B is also substantial,

Table 2a
Results from 2-way ANOVA with independent factors Voice and Place of Articulation for IP-medial stops

Speaker F3A					Speaker F2B						
	Factor	df	F	Sig.	Partial η^2		Factor	df	F	Sig.	Partial η^2
VOT (ms)	VOICE	1, 181	31.448	.000	.148	VOT	VOICE	1, 205	186.454	.000	.476
	POA	2, 181	4.047	.019	.043		POA	2, 205	10.448	.000	.093
	VOI*POA	2, 181	.373	.689	.004		VOI*POA	2, 205	1.893	.153	.018
f_0 (Hz)	VOICE	1, 275	31.291	.000	.102	f_0	VOICE	1, 234	4.310	.039	.018
	POA	2, 275	.629	.534	.005		POA	2, 234	1.549	.215[†]	.013
	VOI*POA	2, 275	7.426	.001	.051		VOI*POA	2, 234	1.996	.138	.017
CD (ms)	VOICE	1, 211	9.587	.002	.043	CD	VOICE	1, 169	.623	.431	.004
	POA	2, 211	7.641	.001	.068		POA	2, 169	15.213	.000	.153
	VOI*POA	2, 211	2.060	.130	.019		VOI*POA	2, 169	2.936	.056	.034
BA (dB)	VOICE	1, 273	2.306	.130	.008	BA	VOICE	1, 240	4.508	.035	.018
	POA	2, 273	73.465	.000	.350		POA	2, 240	51.805	.000	.302
	VOI*POA	2, 273	8.763	.000	.060		VOI*POA	2, 240	6.589	.002	.052

Significant results in bold. Speakers F3A and F2B.

[†]Indicates that the result from the Kruskal–Wallis test shows different significance.

Table 2b
Results from 2-way ANOVA with independent factors Voice and Place of Articulation for IP-medial stops

Speaker M1B					Speaker M2B						
	Factor	df	F	Sig.	Partial η^2		Factor	df	F	Sig.	Partial η^2
VOT (ms)	VOICE	1, 249	209.942	.000	.457	VOT	VOICE	1, 156	54.990	.000	.261
	POA	2, 249	16.454	.000	.117		POA	2, 156	4.125	.018	.050
	VOI*POA	2, 249	7.937	.000	.060		VOI*POA	2, 156	.122	.885	.002
f_0 (Hz)	VOICE	1, 294	19.252	.000	.061	f_0	VOICE	1, 227	10.447	.001	.044
	POA	2, 294	.380	.684	.003		POA	2, 227	2.450	.089 [†]	.021
	VOI*POA	2, 294	6.827	.001	.044		VOI*POA	2, 227	.545	.581	.005
CD (ms)	VOICE	1, 220	.048	.827	.000	CD	VOICE	1, 174	5.163	.024[†]	.029
	POA	2, 220	41.688	.000	.275		POA	2, 174	28.270	.000	.245
	VOI*POA	2, 220	1.478	.230	.013		VOI*POA	2, 174	2.903	.058	.032
BA (dB)	VOICE	1, 224	2.404	.122	.008	BA	VOICE	1, 226	52.919	.000	.190
	POA	2, 224	5.974	.003	.39		POA	2, 226	10.148	.000	.082
	VOI*POA	2, 224	.200	.819	.001		VOI*POA	2, 226	2.666	.072	.023

Significant results in bold. Speakers M1B and M2B.

[†]Indicates that the result from the Kruskal–Wallis test shows different significance.

accounting for 19% of BA variance, but all other Voice effects remain small, at or under 10%. All speakers show significant Voice effects on at least two acoustic measures, and speaker M2B shows significant Voice effects on all four acoustic measures. When the results from 2-way ANOVA are compared with the results from the Kruskal–Wallis test for Voice groups that show unequal variance under Levene's test, only one difference emerges. The significant effect from 2-way ANOVA for speaker M2B on the acoustic measure of CD fails to reach significance under Kruskal–Wallis ($X^2(1,178) = 3.737$, sig. = .053).

Significant effects of POA are also observed in Tables 2a,b for all speakers for three of the four acoustic measures. No POA effect was observed on f_0 . Speakers differ quite a bit in the size of the POA effect on each acoustic measure. The POA effect with the greatest η^2 value is BA for speakers F3A (35%) and F2B (30.2%),

Table 3
Results of post hoc analysis of POA grouping for IP-medial stops, for each speaker individually

Speaker	Dependent measure	POA grouping
F3A	VOT	Labial = Alveolar < Velar
	f ₀	Labial < Velar; Alveolar = Labial; Alveolar = Velar
	CD	Alveolar = Velar < Labial
	BA	Labial = Velar < Alveolar
F2B	VOT	Labial < Alveolar < Velar
	f ₀	Labial < Velar; Alveolar = Labial; Alveolar = Velar
	CD	Alveolar = Velar < Labial
	BA	Labial = Velar < Alveolar
M1B	VOT	Labial < Alveolar < Velar
	f ₀	Labial = Velar = Alveolar
	CD	Alveolar = Velar < Labial
	BA	Labial < Alveolar; Velar = Labial; Velar = Alveolar
M2B	VOT	Labial < Alveolar = Velar
	f ₀	Labial < Velar; Alveolar = Labial; Alveolar = Velar
	CD	Alveolar = Velar < Labial
	BA	Alveolar < Velar < Labial

The “<” symbol marks significant differences at the .05 level under both the Tukey HSD and the nonparametric Tamhane tests. The “=” symbol indicates that the two groups are not significantly different under either test.

and CD for speakers M1B (27.5%) and M2B (24.5%), but there are additional POA effects over 10% (F2B: 15.3% for CD; M1B: 11.7% for VOT). Comparison of the 2-way ANOVA results with the results from the Kruskal–Wallis tests, for those groups that have unequal variance by Levene’s test, shows that all of the significant effects from ANOVA are also significant under Kruskal–Wallis analysis. Additional findings from the Kruskal–Wallis test are significant POA effects on f₀ for two speakers (F2B: $X^2(2,238) = 11.233$; $p < .005$; M2B: $X^2(2,233) = 6.984$; $p < .05$). As shown in Table 3, post-hoc tests (both Tukey–HSD and Tamhane) show significant differences at the .05 level between the POA groups for VOT: Labial < Alveolar (3 speakers) and Alveolar < Velar (3 speakers); for f₀: Labial < Velar (3 speakers), Labial = Alveolar (all speakers), Alveolar = Velar (all speakers); and for CD: Alveolar = Velar < Labial (all speakers). The POA grouping for BA is also significant, but variable across speakers. The main trend is Labial < Alveolar, with Velar either nondistinct from Labial (for speakers F3A, F2B), or from both Labial and Alveolar (for speaker M1B). Speaker M2B shows an unusual pattern where Labial has the highest BA value: Alveolar < Velar < Labial.

There are relatively few significant interactions between Voice and POA, and no consistent patterns of interaction across speakers. Speaker M2B shows no interactions. For the other speakers, there are interactions for the dependent measures of VOT (M1B), f₀ (F3A, M1B), and BA (F3A, F2B), but these are all small effects based on η^2 values, accounting for no more than 6% of the variance for the relevant acoustic measure. In all cases of a Voice \times POA interaction, there is also a main effect for either Voice or POA on the acoustic measure that is much bigger in its effect size.

A significant effect of POA on burst spectrum measures is observed on each acoustic measure for each speaker in the 1-way ANOVA results in Table 4 (descriptive statistics are shown in Table A4a,b of the Appendix). The only difference between speakers is in the strength of the effect on the BS–mid measure, which is significant at the level of $p < .001$ for all speakers except M2B, where the effect is only marginally significant at the level of $p < .05$. The biggest effect of POA is on the BS–hi measure for each speaker, where POA accounts for between 50% (for M1B and M2B) and 72% (for F3A) of the variance. This finding indicates that POA has the greatest influence on energy in the upper region (3–5 kHz) of the frequency range. The effect of POA on BS–mid (in the 1–3 kHz range) is also substantial for three of the speakers (F3A, F2B, M1B), accounting for around 15–40% of the overall variance in that measure. POA effects on BS–low (in the 0–1 kHz range) are much smaller for all speakers, as is the effect on BS–mid for speaker M2B, accounting for

Table 4

Results of one-way ANOVA for the effects of POA on burst spectrum measures for IP-medial stops, for each speaker individually

Speaker	Burst spectrum measure (dB)	F	df	Sig.	Partial η^2
F3A	BS–low	10.305	2, 278	.000	.069
	BS–mid	91.014	2, 278	.000	.396
	BS–hi	364.345	2, 278	.000	.724
	BSRatio–low	89.945	2, 278	.000	.393
	BSRatio–hi	177.004	2, 278	.000	.560
F2B	BS–low	11.476	2, 243	.000	.086
	BS–mid	21.882	2, 243	.000	.153
	BS–hi	174.030	2, 243	.000	.589
	BSRatio–low	50.523	2, 243	.000	.294
	BSRatio–hi	107.691	2, 243	.000	.470
M1B	BS–low	11.431	2, 300	.000	.071
	BS–mid	31.886	2, 300	.000	.175
	BS–hi	150.922	2, 300	.000	.502
	BSRatio–low	65.899	2, 300	.000	.305
	BSRatio–hi	192.769	2, 300	.000	.562
M2B	BS–low	12.976	2, 229	.000	.102
	BS–mid	3.263	2, 229	.040[†]	.028
	BS–hi	115.403	2, 229	.000	.502
	BSRatio–low	5.243	2, 229	.006	.044
	BSRatio–hi	109.646	2, 229	.000	.489

Significant results in bold. Partial η^2 values estimate effect size.[†]Indicates that the result from the Kruskal–Wallis test has different significance.

around 10% or less of the overall variance.³ The Kruskal–Wallis test was also applied on the basis of results from Levene's test showing unequal variance for many of the comparison groups, and showed highly significant effects of POA on all acoustic measures for all speakers with only one exception: BS–mid for speaker M2B. This measure was also the only one that was marginal under the ANOVA analysis.

The results of post-hoc analysis are shown in Table 5. Post-hoc analysis based on Tukey–HSD and Tamhane tests reveals significant differences at the .05 level among the POA groups for BS–low: Alveolar < Labial (3 speakers) and Alveolar = Velar (3 speakers); for BS–mid: Labial < Alveolar (3 speakers), with Velar either non-distinct from Alveolar (2 speakers) or non-distinct from both Alveolar and Labial (1 speaker); for BS–hi: Labial < Velar < Alveolar (all speakers); for BSRatio–low: Labial < Alveolar (all speakers), with Velar either non-distinct (3 speakers) or distinct (1 speaker) from Alveolar; for BSRatio–hi: Alveolar < Velar (all speakers), with Labial either non-distinct (2 speakers) or distinct (2 speakers) from Velar.

3.3.2. Effects of Accent on VOT, f_0 , CD and BA in IP-medial stops

The results of 2-way ANOVA with the independent factors of Accent and Voice are shown in Tables 6a,b. A significant effect of Accent is found on VOT and CD for all speakers. Three speakers (F3A, F2B, M1B) show a significant effect of Accent on f_0 , and only speaker M1B shows a significant effect on BA. η^2 values for the Accent effect on four acoustic measures are under 10% for all speakers, except CD for speaker F3A (16%). A comparison of η^2 values between Accent and Voice reveals, not surprisingly, that for the VOT measure Voice accounts for a greater portion of the observed variation than Accent does, for all speakers. The Voice effect on f_0 is also greater than the Accent effect for all speakers except speaker F2B (Accent: 5.8%;

³A reviewer suggests that the small effect size of POA on BS–low may be due to the fact that voicing is a primary source of variability on BS–low in the burst release. This hypothesis was tested and confirmed for speaker F3A through ANOVA with BS–low as the dependent variable and the independent factors Voice and POA. While both factors are significant at $p < .01$, the η^2 values show a larger effect of Voice on BS–low variation (.121) than of POA (.018).

Table 5
Results of post hoc analysis of POA grouping for IP-medial stops based on burst spectrum measures

Speaker	Dependent measure (dB)	POA grouping
F3A	BS–low	Velar < Labial = Alveolar
	BS–mid	Labial < Alveolar = Velar
	BS–hi	Labial < Velar < Alveolar
	BSRatio–low	Labial < Alveolar = Velar
	BSRatio–hi	Alveolar < Velar < Labial
F2B	BS–low	Velar = Alveolar < Labial
	BS–mid	Labial < Alveolar = Velar
	BS–hi	Labial < Velar < Alveolar
	BSRatio–low	Labial < Alveolar = Velar
	BSRatio–hi	Alveolar < Velar = Labial
M1B	BS–low	Alveolar < Labial; Labial = Velar; Alveolar = Velar
	BS–mid	Labial < Alveolar < Velar
	BS–hi	Labial < Velar < Alveolar
	BSRatio–low	Labial < Alveolar < Velar
	BSRatio–hi	Alveolar < Velar = Labial
M2B	BS–low	Velar = Alveolar < Labial
	BS–mid	Labial = Alveolar = Velar
	BS–hi	Labial < Velar < Alveolar
	BSRatio–low	Labial < Alveolar; Labial = Velar; Alveolar = Velar
	BSRatio–hi	Alveolar < Velar < Labial

Voice: 5.5%). The Voice effects on CD and BA are smaller than the Accent effect for all speakers except BA for speakers F2B and M2B, but most of the Voice and Accent effects remain substantially small, under 10%.

For Accent groups that show unequal variance under Levene's test, the 2-way ANOVA results are compared with results from the Kruskal–Wallis test. Most of the observed effects are significant under both tests, except VOT in speaker F2B: the significant effect from 2-way ANOVA for VOT in speaker F2B fails to reach significance under Kruskal–Wallis: $X^2(1207) = 2.587$, sig. = .108).

There is a significant interaction between Accent and Voice on VOT for all speakers. Three speakers (F3A, F2B, M1B) show a significant interaction on f_0 , two speakers (M1B, M2B) show a significant interaction on CD, and two speakers (F3A, M2B) show a significant interaction on BA.

In summary, acoustic measures of VOT, f_0 and CD show significant effects of Accent for most of the speakers, with interactions between Accent and Voice primarily for VOT and f_0 . Interaction graphs in Fig. 2 illustrate the direction and degree of change under accent for VOT, f_0 and CD as voicing cues. Accent conditions are listed on the x-axis and the means of the acoustic measure are plotted on the y-axis. It is shown that Accent has a raising effect on VOT, f_0 and CD for both voiced and voiceless stops for all speakers, except for VOT in voiced stops for speaker F2B and M1B. In addition, Accent has a greater effect of increase on voiceless stops for VOT and f_0 , and on voiced stops for CD, except f_0 for M2B and CD for M1B. These results suggest that while the general effect of Accent is to increase VOT, f_0 and CD for all stops, the degree of increase is often unequal for voiced and voiceless stops with the result that the voiced and voiceless categories are frequently more separated along these three acoustic dimensions under Accent.

In order to examine the effect of Accent on POA cues, 2-way ANOVA was performed for VOT, CD and BA, but not for f_0 since in Section 3.3.1 f_0 is found not to encode POA categories for any speaker. The results of 2-way ANOVA with the independent factors of Accent and POA are shown in Table 7. A significant effect of Accent is observed on VOT and CD for all speakers. Speaker M1B shows a significant effect of Accent on BA. η^2 values for the Accent effect remain small, near or under 10% for all speakers, except VOT for speaker F3A (24.4%). A comparison of η^2 values between Accent and POA reveals that POA accounts for a greater

Table 6a
Two-way ANOVA with independent factors Accent and Voice for IP-medial stops

Speaker F3A						Speaker F2B					
	Factor	df	F	Sig.	η^2		Factor	df	F	Sig.	η^2
VOT (ms)	ACCENT	1, 183	11.625	.001	.060	VOT	ACCENT	1, 207	11.684	.001 [†]	.053
	VOICE	1, 183	37.974	.000	.172		VOICE	1, 207	291.070	.000	.584
	ACC*VOI	1, 183	4.737	.031	.025		ACC*VOI	1, 207	30.424	.000	.128
f ₀ (Hz)	ACCENT	1, 277	17.240	.000	.059	f ₀	ACCENT	1, 236	14.456	.000	.058
	VOICE	1, 277	43.530	.000	.136		VOICE	1, 236	13.843	.000	.055
	ACC*VOI	1, 277	4.806	.029	.017		ACC*VOI	1, 236	5.906	.016	.024
CD (ms)	ACCENT	1, 213	40.266	.000	.159	CD	ACCENT	1, 171	12.215	.001	.067
	VOICE	1, 213	38.145	.000	.152		VOICE	1, 171	5.797	.017	.033
	ACC*VOI	1, 213	1.560	.213	.007		ACC*VOI	1, 171	2.301	.131	.013
BA (dB)	ACCENT	1, 275	.411	.522	.001	BA	ACCENT	1, 242	.010	.920	.000
	VOICE	1, 275	.000	.997	.000		VOICE	1, 242	.122	.727	.001
	ACC*VOI	1, 275	4.168	.042	.015		ACC*VOI	1, 242	1.023	.313	.004

Significant results in bold. Speakers F3A and F2B.

[†]Indicates that the result from the Kruskal–Wallis test has different significance.

Table 6b
Two-way ANOVA with independent factors Accent and Voice for IP-medial stops

Speaker M1B						Speaker M2B					
	Factor	df	F	Sig.	η^2		Factor	df	F	Sig.	η^2
VOT (ms)	ACCENT	1,251	13.719	.000	.052	VOT	ACCENT	1,158	8.820	.003	.053
	VOICE	1,251	240.749	.000	.490		VOICE	1,158	85.166	.000	.350
	ACC*VOI	1,251	15.790	.000	.059		ACC*VOI	1,158	6.028	.015	.037
f ₀ (Hz)	ACCENT	1,296	18.071	.000	.058	f ₀	ACCENT	1,229	1.888	.171	.008
	VOICE	1,296	23.408	.000	.073		VOICE	1,229	15.391	.000	.063
	ACC*VOI	1,296	11.322	.001	.037		ACC*VOI	1,229	.066	.797	.000
CD (ms)	ACCENT	1,222	22.246	.000	.091	CD	ACCENT	1,176	18.643	.000	.096
	VOICE	1,222	9.090	.003	.039		VOICE	1,176	6.316	.013	.035
	ACC*VOI	1,222	4.738	.031	.021		ACC*VOI	1,176	8.907	.003	.048
BA (dB)	ACCENT	1,296	9.570	.002	.031	BA	ACCENT	1,228	.668	.415	.003
	VOICE	1,296	.174	.677	.001		VOICE	1,228	68.241	.000	.230
	ACC*VOI	1,296	.246	.620	.001		ACC*VOI	1,228	6.025	.015	.026

Significant results in bold. Speakers M1B and M2B.

portion of the variation than Accent does for most cases. The Accent effect is bigger than the POA effect only in sporadic cases such as VOT for speaker F3A and BA for speaker M1B.

There are a few sporadic significant interactions between Accent and POA. Speakers F2B and M1B show a significant interaction between Accent and POA for VOT, while an interaction effect for BA is found only with speaker F3A.

In summary, the mean acoustic measures of VOT and CD as cues to POA show significant effects of Accent for all speakers. Interaction graphs in Fig. 3 show the direction and degree of change under accent for each POA category. Accent conditions and the means of the acoustic measure are plotted on the x-axis and y-axis, respectively. Although significant interactions are present in only a few cases, the Accent effect on VOT is significantly different for labials than for velars and alveolars for two speakers, F2B and M1B, with a similar

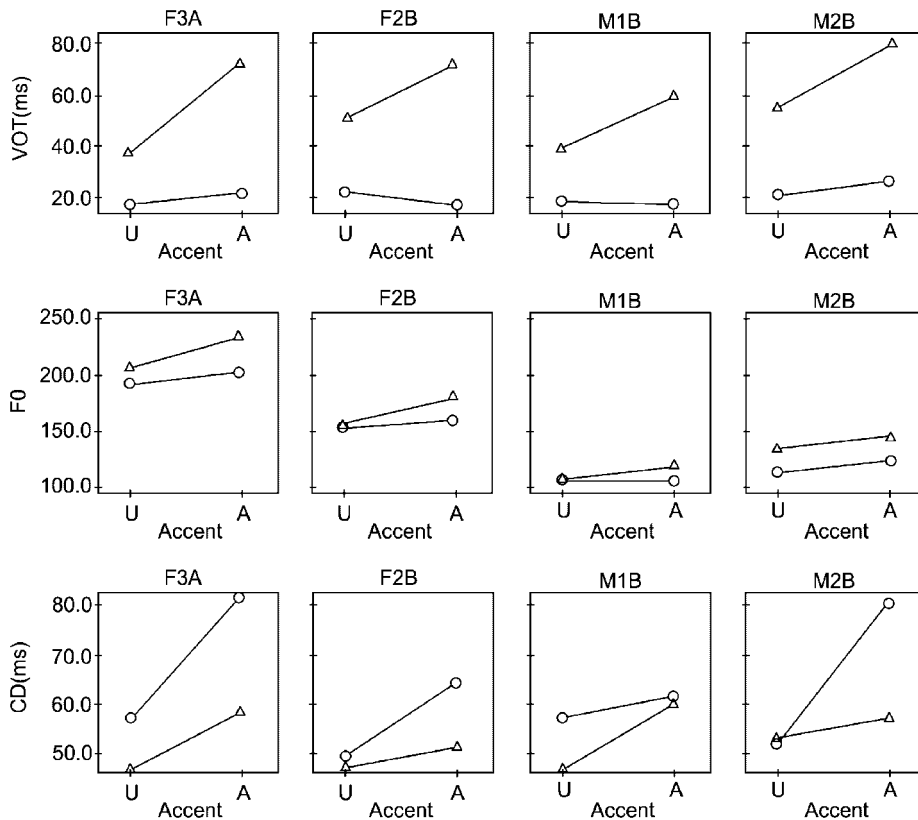


Fig. 2. Interaction graphs illustrating the effect of Accent on VOT, F0, and CD as voicing cues for four speakers in Unaccented (U) and Accented (A) positions: O, Voiced; Δ, Voiceless.

though non-significant trend for M2B. The result of this interaction is that the place distinction between labials and velars/alveolars is greater in accented tokens compared to unaccented tokens. There is no comparable enhancement of the place distinction between velars and alveolars. The Accent effect on CD, on the other hand, is more uniform across POA categories. Accordingly, the Accent effect contributes in only a very limited way to a greater separation of POA categories in acoustic space.

As a measure of the effect of Accent on acoustic variability, standard deviations were compared for the acoustic measures VOT, f_0 , CD and BA in Accented and Unaccented conditions. No consistent pattern of difference was found. For speaker F3A standard deviations for all four measures are greater in the Accented condition ($sd_A > sd_U$); F2B shows $sd_A > sd_U$ for VOT and CD, and $sd_U > sd_A$ for f_0 and BA; M1B shows $sd_A > sd_U$ for VOT, f_0 , and BA, and $sd_U > sd_A$ for CD; M2B shows $sd_A > sd_U$ for VOT and CD, $sd_U > sd_A$ for BA, and $sd_A = sd_U$ for f_0 .

3.3.3. The interaction between lexical stress and phrasal accent

Only lexically stressed syllables are eligible to receive phrasal accent in English. That means that the stops in the Accented condition are also in stressed syllables, whereas the Unaccented condition may include stops from both stressed and unstressed syllables. This situation calls into question the extent to which the Accent effects reported above might be attributed to lexical stress. To determine whether Accent effects are independent of stress effects on the acoustic measures of VOT, f_0 , CD, and BA, statistical analyses were run again for data from one speaker, F3A, this time making a further distinction between stressed and unstressed stops in the Unaccented condition. Each stop token in speaker F3A's data was coded for lexical stress on the basis of the primary stress indicated in the dictionary entry for each word. Out of 185 Unaccented stops (in IP-medial position) 135 were unstressed, and 50 were stressed. For comparison, there were 96 stops in the

Table 7

Two-way ANOVA with independent factors Accent and POA for IP-medial stops, for each speaker individually

Speaker F3A					Speaker F2B						
	Factor	df	F	Sig.	Partial η^2		Factor	df	F	Sig.	Partial η^2
VOT (ms)	ACCENT	1,181	58.442	.000	.244	VOT	ACCENT	1,205	9.813	.002	.046
	POA	2,181	14.129	.000	.135		POA	2,205	28.946	.000	.220
	ACC*POA	2,181	2.947	.055	.032		ACC*POA	2,205	8.598	.000	.077
CD (ms)	ACCENT	1,211	27.343	.000	.115	CD	ACCENT	1,169	11.430	.001	.063
	POA	2,211	15.169	.000	.126		POA	2,269	16.605	.000	.164
	ACC*POA	2,211	2.083	.127	.019		ACC*POA	2,269	.612	.544	.007
BA (dB)	ACCENT	1,273	.793	.374	.003	BA	ACCENT	1,240	.303	.583	.001
	POA	2,273	81.480	.000	.374		POA	2,240	46.751	.000	.280
	ACC*POA	2,273	4.391	.013	.031		ACC*POA	2,240	1.276	.281	.011
Speaker M1B					Speaker M2B						
VOT (ms)	ACCENT	1,249	30.625	.000	.110	VOT	ACCENT	1,156	16.318	.000	.095
	POA	2,249	39.811	.000	.242		POA	2,256	12.327	.000	.136
	ACC*POA	2,249	6.440	.002	.249		ACC*POA	2,256	2.270	.107	.028
CD (ms)	ACCENT	1,220	25.568	.000	.104	CD	ACCENT	1,174	14.888	.000	.079
	POA	2,220	45.956	.000	.295		POA	2,174	29.651	.000	.254
	ACC*POA	2,220	1.205	.302	.011		ACC*POA	2,174	.392	.676	.004
BA (dB)	ACCENT	1,294	13.690	.000	.044	BA	ACCENT	1,226	1.102	.295	.005
	POA	2,294	5.513	.004	.036		POA	2,226	15.717	.000	.122
	ACC*POA	2,294	1.517	.221	.010		ACC*POA	2,226	1.455	.236	.013

Significant results in bold.

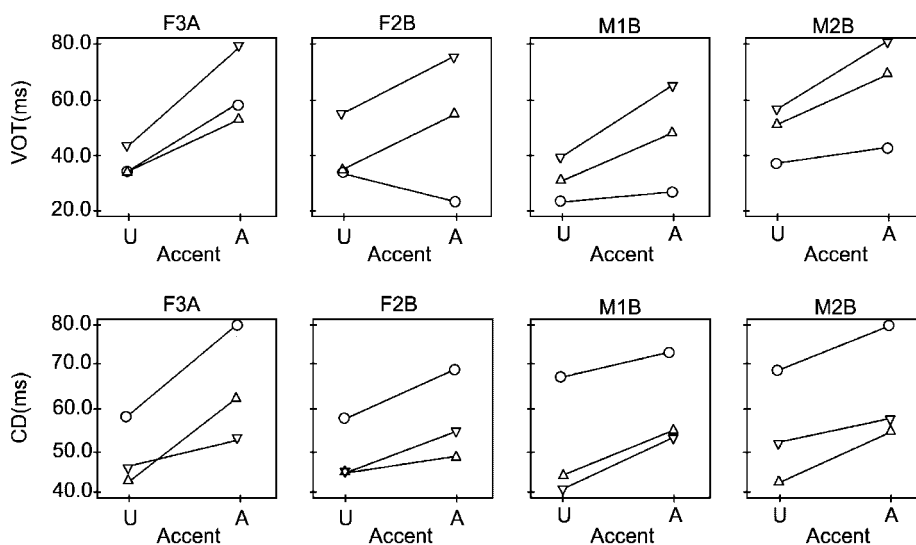


Fig. 3. Interaction graphs illustrating the effect of Accent on VOT and CD as POA cues, for four speakers in Unaccented (U) and Accented (A) positions: ○, labial; Δ, alveolar; ∇, velar.

Accented group for the same speaker (all of which were stressed). Two-way ANOVA was performed with the independent factors Stress/Accent and Voice and the dependent variables VOT, f_0 , CD, and BA. The Stress/Accent factor distinguishes three groups: stressed–accented, stressed–unaccented, and unstressed–unaccented,

and the Voice factor distinguishes voiced from voiceless stops. Results show significant effects of Stress/Accent on VOT ($F(2, 132) = 3.827, p < .05$) and f_0 ($F(2,132) = 4.373, p < .05$), and no significant effect on BA. These findings are consistent with the results reported above for the simpler Accent condition that compared Accented stops with pooled stressed and unstressed Unaccented stops. However, unlike the earlier findings, there was no significant effect of Stress/Accent on CD. Another 2-way ANOVA was performed with the independent factors Stress/Accent and POA (Labial, Alveolar, Velar), and the results show significant effects of Stress/Accent on VOT ($F(2, 132) = 26.984, p < .001$) and CD ($F(2,132) = 6.768, p < .001$), and no significant effect on BA. These findings are identical to those for the simpler Accent condition. In addition, there was a significant effect of Stress/Accent on f_0 ($F(2, 132) = 10.521, p < .001$), an effect which was not observed under the simple Accent condition. There were no significant interactions of Stress/Accent with either Voice or POA.

Post-hoc analysis of the effects of Stress/Accent are revealing. When stops are grouped for Voice and Stress/Accent, the results from Tukey's test groups the unstressed–unaccented and stressed–unaccented stops together for the dependent measures of f_0 and CD, and maintains a 3-way distinction between the Stress/Accent groups for the VOT measure. The same pattern of results holds for Tukey's test when stops are grouped for POA and Stress/Accent. Importantly, there are no results that suggest a grouping of stressed–accented and stressed–unaccented stops together, providing clear evidence that Accent effects on these measures go beyond the effects that can be attributed to lexical stress alone. This finding supports the exclusion of stress as an independent factor in the analysis of Accent effects on VOT, f_0 , CD and BA presented above, and on burst spectral measures presented in Section 3.3.4.

3.3.4. Effects of Accent on burst spectra for IP-medial stops

Results of 2-way ANOVA for the effect of Accent on burst spectrum measures show that there is a significant effect of Accent only in sporadic cases (BS–low for speakers F3A and M1B; BSRatio–low for speakers F3A, F2B and M1B). In addition, there are only few significant interactions of Accent and POA across speakers (BS–low for speakers F3A and M1B; BSRatio–low for speaker M1B. Table 8 shows results of 2-way ANOVA only for BS–low and BSRatio–low, since no significant effects of Accent are found for the other burst spectrum measures. η^2 values for the Accent effect remain small, under 6% for all speakers. A comparison of η^2 values between Accent and POA shows that POA accounts for a much greater portion of the variation than Accent does for all cases.

For the cases where significant effects of Accent are found (BS–low for speakers F3A and M1B; BSRatio–low for speakers F3A, F2B and M1B), interaction graphs are plotted to show the direction and degree of change under accent for each POA category, as shown in Fig. 4. Fig. 4 also shows that for BS–low, Accent has a lowering effect for labials, but a raising effect for alveolars and velars. For BSRatio–low, Accent has a lowering effect for all POAs. Overall, the acoustic distance between POA categories under the Accented condition, as measured by the burst spectrum, is not greater than under the Unaccented condition.

A comparison of standard deviations for burst spectrum measures in Accented and Unaccented conditions shows no consistent pattern of difference within or across speakers. Summarizing over all five Burst Spectrum measures for four speakers (20 cases), standard deviations are greater in the Accented condition in 11 cases ($sd_A > sd_U$), while $sd_U < sd_A$ in six cases, and $sd_A = sd_U$ in two cases.

3.3.5. Effects of IP-position on acoustic measures of Voicing for /t,d/

In order to examine the effect of IP-position on the acoustic cues to voicing for /t/ and /d/, analysis was done first for only unaccented /t, d/. Based on the results from Levene's test for equality of variance, 2-way ANOVA was performed with the independent factors of Position and Voice. There is no significant effect of Position on any of the acoustic measures (VOT, f_0 , BA or CD) for any speakers. In addition, in only a few cases is there a significant interaction between Position and Voice. Another 2-way ANOVA was performed for the combined accented, unaccented /t,d/ data, but again no significant effects of IP-position or interactions between IP-position and Voice are observed.

As a measure of variance in IP-initial and IP-medial conditions, standard deviations are compared for each acoustic measure. In general, IP-initial /t/ and /d/ have smaller standard deviations than IP-medial /t, d/,

Table 8
Two-way ANOVA results for effects of ACCENT on burst spectrum measures

Speaker	Dependent measure (dB)	Factor	F	df	Sig.	Partial η^2
F3A	BS-low	ACCENT	4.777	1, 275	.030	.017
		POA	12.042	2, 275	.000	.081
		ACCENT*POA	5.810	2, 275	.003	.041
	BSRatio-low	ACCENT	5.871	1, 275	.016	.021
		POA	77.511	2, 275	.000	.360
		ACCENT*POA	1.192	2, 275	.305	.009
F2B	BS-low	ACCENT	1.822	1, 240	.178	.008
		POA	10.040	2, 240	.000	.077
		ACCENT*POA	1.135	2, 240	.323	.009
	BSRatio-low	ACCENT	4.813	1, 240	.029	.020
		POA	36.682	2, 240	.000	.234
		ACCENT*POA	1.147	2, 240	.319	.009
M1B	BS-low	ACCENT	6.161	1, 297	.014	.020
		POA	7.487	2, 297	.001	.048
		ACCENT*POA	3.529	2, 297	.031	.023
	BSRatio-low	ACCENT	16.378	1, 297	.000	.052
		POA	62.329	2, 297	.000	.296
		ACCENT*POA	4.909	2, 297	.008	.032
M2B	BS-low	ACCENT	2.500	1, 226	.115	.011
		POA	11.156	2, 226	.000	.090
		ACCENT*POA	2.954	2, 226	.054	.025
	BSRatio-low	ACCENT	1.672	1, 226	.197	.007
		POA	4.771	2, 226	.009	.041
		ACCENT*POA	1.204	2, 226	.302	.011

Results are shown only for those burst spectrum measures that showed a main effect of Accent.

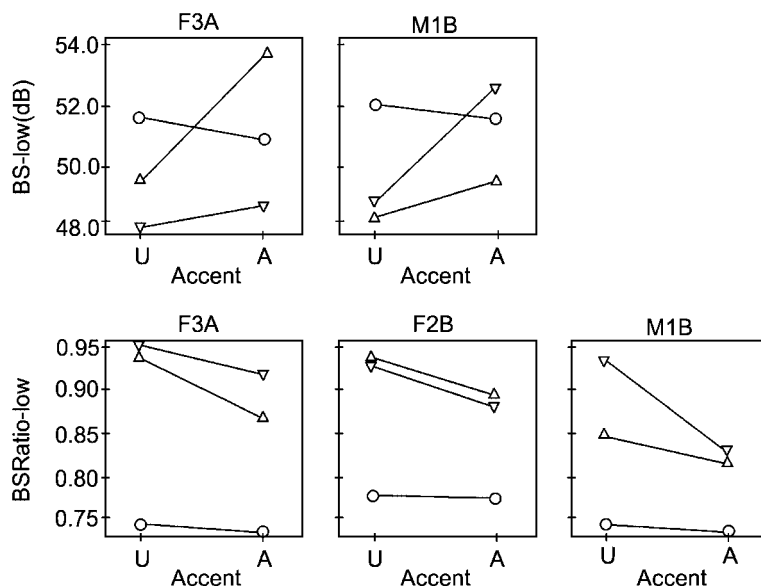


Fig. 4. Interaction graphs measure illustrating the effect of Accent on mean burst spectrum measures as POA cues, in Unaccented (U) and Accented (A) positions. ○, labial; △, alveolar; ▽, velar.

suggesting that the articulation in IP-initial position is more precisely controlled by the speaker than in IP-medial positions.

4. Discussion

4.1. Acoustic correlates of Voice and POA

Our first finding in this study is that significant correlates of the Voice and POA distinctions are found in the acoustic measures of VOT, f_0 , CD, BA and burst spectrum. Speakers are indeed using these acoustic properties to encode the phonological contrasts of Voice and POA. The specific patterns of our results partially confirm those of the prior studies cited above (Section 2, Hypotheses 1 and 2) for acoustic correlates of Voice and POA. The encoding of Voice distinctions in VOT and f_0 is consistent across studies: voiceless stops have higher VOT and f_0 values than voiced stops. The pattern of CD and BA as Voice cues in our data differ from prior findings: in the data presented here CD is significantly longer in voiced stops than in voiceless for three out of four speakers, and BA is significantly lower in voiceless stops than in voiced for three speakers. While voicing effects on CD are not consistently found in earlier studies, when observed they are always in the opposite direction of our findings, with shorter CD for voiced stops. Since the speech materials are not uniform across studies, it is not possible to know which factors are responsible for the conflicting findings, although we suspect that the local stress and syllable context, which are not controlled in our study, may play a role. The pattern of BA results in our data are also puzzling, although there is little opportunity for comparison with other studies as most studies do not report on BA as a cue to stop voicing. Further research on CD and BA as Voice cues is clearly required.

As for the encoding of POA, the current results are largely consistent with prior findings for the VOT measure. We find Labial < Alveolar for three out of four speakers, and Alveolar < Velar, also for three out of four speakers. It is worth noting that the long VOT for velars in our study is consistent with the findings of [Cho and Ladefoged \(1999\)](#) for 18 languages. The results of CD as a cue to POA partially confirm prior findings, primarily in demonstrating that labials have longer CD than velars and alveolars, for all four speakers. BA values are significantly different by POA for all four speakers, though the patterns vary across speakers. The main pattern is Labial < Alveolar, which coincides with previous studies. Speaker M2B demonstrates an unusual pattern with Alveolar < Velar < Labial. The results for burst spectrum measures also confirm prior findings, demonstrating that the spectral envelope of the stop release burst plays a role in encoding the POA contrasts.

Overall, we observe considerable variation across speakers in the way that Voice and POA contrasts are expressed through the acoustic properties measured here. This is not in itself a surprising result, and would be expected to follow from inter-speaker variation in articulation. Even small differences in laryngeal and supralaryngeal articulation across speakers would affect many or all of the acoustic measures, which are physiologically interdependent and which collectively express the phonological category. As for voicing cues, M2B shows significant voice effects on all acoustic measures while the other speakers show significant effects only on some of the measures (F3A: VOT, f_0 and CD; F2B: VOT, f_0 and BA; M1B: VOT and f_0). VOT and f_0 are significant Voice cues for all speakers, but a comparison of η^2 values across the acoustic measures for each speaker reveals that Voice accounts for a greater portion of the variation in VOT, at between 15% and 48% across speakers, than it accounts for variation in other acoustic measures. The relative strength of the VOT cue to Voicing holds for all speakers. Of the POA cues, all speakers show significant POA effects on VOT, CD, BA and burst spectrum measures, and post-hoc analyses shows that all three POA categories are distinguished. However, examination of η^2 values shows that the POA effects on the burst spectrum measures, and on BS–hi in particular, are greater than the POA effects on the other acoustic measures. POA accounts for between 50% and 72% of the variation in BS–hi across speakers, while it accounts for only 7–28% of the variation in CD and 4–35% of the variation in BA, and even less of the variation in VOT and f_0 . From these results we conclude that POA is most effectively cued through variation in the burst spectrum measures in these data.

One notable feature of our data from radio news speech is the insufficiency of any individual acoustic property in classifying stops for either Voice or POA. For instance, although Voice conditions more variation

with VOT than with any other acoustic property, for all speakers, the voiced and voiceless stops are not well separated into two groups based on VOT values alone. Rather, the distribution of VOT values of voiced and voiceless stops are overlapping to some degree for labials, alveolars and velars, for every speaker. A similar situation obtains for the separation of POA categories based on VOT, CD, BA or burst spectrum measures. No individual acoustic property serves to uniquely identify the POA category of a stop; for each acoustic measure the distributions of values for stop tokens grouped by POA are overlapping to varying degrees, for every speaker.

This is not to say that the acoustic properties share equally in marking the Voice or POA contrasts. The η^2 values reported above indicate that VOT is the measure that is the most affected by the Voice contrast and that burst spectrum measures, especially BS–hi, are the measures most affected by the POA contrast. Beyond the η^2 values, there are also observable differences between acoustic measures in the degree to which they serve to separate the Voice or POA categories. For example, voiced and voiceless stops are generally more separated along the VOT dimension than they are along the dimensions defined by f_0 , CD or BA. Similarly, POA categories are somewhat more separated along the BS–hi dimension than along dimensions defined by BS–low, or by VOT, CD or BA. Based on such differences, we could say that VOT is a **primary** acoustic marker of the Voice contrast, and BS–hi is a **primary** acoustic marker for the expression of POA. However, this designation of an acoustic measure as primary in marking a phonological contrast must be tempered by the recognition that no acoustic measure effectively classifies stop tokens on its own, and in some cases the degree of separability between acoustic measures differs only slightly. Furthermore, the separability of tokens by Voice or by POA along any acoustic dimension varies somewhat by speaker, and there are also interactions between Voice and POA in category separation. Given the extent of overlap for Voice or POA categories along all acoustic dimensions studied here, it seems that for these data multiple acoustic cues must jointly contribute to marking the phonological Voice and POA contrasts, and it is likely that listeners integrate information from multiple acoustic cues in identifying the Voice or POA category of a given stop token. Perception studies would be required to determine the relative contribution of each acoustic measure (i.e., its contrastive status) to the identification of stops according to Voice or POA categories.

4.2. Effects of Accent on the acoustic correlates of Voice and POA

Our second important finding is that Accent has a significant effect on the mean values of VOT, f_0 , and CD as acoustic correlates of Voice, with the biggest effect on CD, where Accent accounts for between 7% and 16% of the CD variation across speakers based on η^2 values. Accent effects for the remaining acoustic measures stand lower, accounting for between 0% and 6% of the variation in those measures across speakers. In addition, a significant interaction between Accent and Voice is found for VOT with all speakers, and for f_0 , BA and CD for some speakers. Overall, we find that Accent has a greater effect on the acoustic measures of voiceless stops for most speakers.

Accent also has an effect on POA cues. As noted above, many of the same acoustic measures that cue Voice also serve to cue POA contrasts in these data, namely VOT, CD and BA. When the data are grouped by Accent and POA factors, ANOVA results indicate again that Accent has significant effects, primarily on VOT and CD. The strength of the Accent effect under this grouping varies by acoustic measure. Based on η^2 values, Accent accounts for between 6% and 12% of the variation in CD across speakers, and accounts for up to 24% of the variation in VOT. Accent also exhibits only a very small effect on the Burst Spectrum measures of POA, accounting for between 1% and 5% of the overall variation in those measures across speakers. Above we observed that Accent affects voiced and voiceless stops differently in many cases, but when the stops are grouped by POA, the Accent effects do not result in overall greater acoustic distinctions between phonologically contrastive POA categories, with the only notable result being the enhanced distinction between labials and non-labials in VOT under accent. For other measures, the effect of Accent appears to be uniform across POA categories, with few significant interactions.⁴ Finally, we note that the effects of Accent

⁴Formant transitions in CV sequences are known to play a primary role as POA cues for stop consonants, and have been shown to exhibit clear effects of prosodic boundary (Tabain, 2003). Though not reported here, we have also examined the effect of Accent on formant transitions as cues to POA for the set of stop data studied here. A preliminary report on the results for speaker F3A appears in

on the individual acoustic cues for Voicing and POA vary considerably in their significance and size across speakers.

In summary, the primary effect of Accent is on variation in CD as a cue to Voicing and on variation in VOT and CD as cues to POA. Lesser effects of Accent (<6%) are observed with the other acoustic measures. The inter-speaker variability in the size of Accent effects on each acoustic measure indicates that there is no uniformity in the expression of Accent through the acoustic correlates of stop Voicing or POA across speakers.

The above findings indicate that the acoustic measures considered here vary systematically both by phonological factors (Voicing and POA) and by the pragmatic factor encoded in the prosodic feature of Accent. To determine which kind of factor has the greater influence, we compare the size of the effects conditioned by each type. A comparison of η^2 values for Voicing and Accent effects shows that for the primary Voice cue, VOT, the Voice effect is far greater than the Accent effect for all speakers. Similarly, f_0 shows a greater effect of Voice than of Accent for three speakers. These two measures are the strongest overall indicators of Voice among the acoustic measures considered here. The lesser effect of Accent on these measures suggests that the phonological function of an acoustic cue takes precedence over its pragmatic function. It is interesting to note that in these data f_0 at the stop release patterns like VOT in that its role in expressing phonological Voicing is as great or greater than its role in expressing pragmatically determined Accent, despite the fact that the vast majority of pitch accents in this corpus have an f_0 peak (marked by a H* in the ToBI labeling scheme) on the accented syllable. In other words, in the short interval following stop release, the microprosodic effect of Voice on f_0 is still evident despite the macroprosodic effect of pitch accent (though both Voice and Accent effects remain small, accounting for less than 8% of the overall variation in f_0 for all but one speaker).

Whereas the Voicing effect dominates the Accent effect for VOT and f_0 , the opposite holds for CD, where the Accent effect is greater. Indeed, the Accent effect is strongest on CD out of all the acoustic measures that encode Voice. CD (along with BA) is a weaker marker of the Voice contrast than VOT or f_0 , across speakers. The general pattern we observe is that Accent effects are strongest for acoustic measures that are weak (non-primary) markers of a phonological contrast, and weakest for acoustic measures that are primary in marking a phonological contrast—a pattern that suggests that the phonological function takes precedence over the pragmatic function in the determination of the acoustic outcome.

Turning next to the comparison of Accent and POA effects based on η^2 values, we find that POA effects dominate with the CD measure for all four speakers, and with the VOT and BA measures for three speakers. POA effects strongly dominate Accent effects for all burst spectrum measures for all speakers. Accent effects dominate POA effects only in two cases: VOT for speaker F3A and BA for speaker M1B; these two acoustic measures were among the weakest measures in encoding the POA contrast. This pattern is similar to the pattern observed in the comparison of Accent and Voice effects: the Accent effects are strongest on the acoustic measures that are weak markers of the phonological POA contrast, and weakest on the primary acoustic measure encoding the POA contrast.

Considering the comparison of Accent effects with both Voice and POA effects, we find that Accent effects are in general contrast preserving, since they have a lesser impact on the acoustic measures that play a primary role in marking a phonological contrast.

(footnote continued)

Cole et al. (2003). Although measurements of F2 at the onset of voicing following stop release were taken for both voiced and voiceless stops, the longer VOT of voiceless stops and Accented stops makes the results difficult to interpret. We observe a small effect of Accent in the correlation between F2-onset and F2 at midpoint of the following vowel, in a locus equation analysis. In plots of F2-onset against F2-midpoint, our findings show regression lines for each POA cluster with slightly steeper slope and lower y-intercept under Accent. This pattern of results suggests greater CV coarticulation under Accent, contrary to expectations, but the interpretation is hindered by the interaction between F2 transition and VOT effects under Accent. In addition, when F2 transition data from all four speakers are considered together, there is no consistent pattern to the Accent effects on F2 transitions across speakers. The pooling of voiced and voiceless stops in our data, the uneven distribution of following vowels, and the confounding effect of long VOT under Accent very likely contribute noise to the F2 measurements. A more reliable study would examine only voiced stops, perhaps with additional data beyond what is available in the corpus studied here.

4.3. Contrast enhancement, uniform strengthening and syntagmatic strengthening effects

The Accent effect on the measures VOT, f_0 , CD and BA indicate a general pattern of prosodic strengthening, seen in the increased values of these measures under Accent for most of the stop consonants. These effects are considered here in light of experimental hypotheses regarding contrast enhancement (Hypotheses 1,2) and syntagmatic strengthening (Hypothesis 3). Considering first the effect of Accent on the acoustic correlates of Voice, the effects are found to be bigger for the voicing group with the higher values for each measure, resulting in greater separation of voicing categories under Accent. In sporadic cases such as VOT in speakers F2B and M1B, VOT values are decreased for voiced stops but increased for voiceless stops. This effect also induces more separation between voiced and voiceless categories. Going against this general trend, there is one case where the effect of Accent actually diminishes the voicing contrast, and that is for the CD measure for speaker M1B (Fig. 2). With the exception of the one case of diminished contrast, these findings define a general pattern of **contrast enhancement** that supports Hypothesis 1: prosodic enhancement of the Voice contrast.

Considering next the effect of Accent on acoustic correlates of POA, for most of the acoustic measures the effects of Accent are found to be more or less uniform across different POA categories, with the result that the degree of separation between POA categories is not greatly changed under Accent—a pattern of **uniform strengthening**. The limited exception to this pattern was the enhanced POA distinction between labials and non-labials in VOT values under accent. In addition, Accent effects on the burst spectrum measures show that the distance between POAs is actually reduced under accent, confirming that Accent does not contribute to the enhancement of contrasting POAs. This finding is contrary to Hypothesis 2, which predicted prosodic enhancement of the POA contrast under Accent. Support for Hypothesis 2 could also come from an observation of reduced variability in the acoustic cues to POA, but there again the findings are contrary. A comparison of standard deviations in Accented and Unaccented conditions for all the acoustic cues to Voice and POA considered in this study reveals no consistent pattern of difference. In some cases variability is actually increased under Accent, though not consistently for any single acoustic measure across speakers. Our findings do not support the claim that Accented sounds are more precise or clearer.

The third hypothesis for this study was that of syntagmatic strengthening: the C/V segmental boundary may be more strongly marked in Accented CVX syllables compared to Unaccented syllables, through longer CD and VOT measures, and higher BA. The results exhibit the predicted strengthening pattern in the durational measures of CD and VOT, but do not demonstrate higher amplitude values for stop bursts in the Accented condition. The longer intervals of closure and VOT should provide stronger auditory cues to the C/V boundary, but in the absence of an effect from BA, no strong conclusions can be drawn. A further limitation is that the present study did not measure prosodic strengthening effects on vowel duration or intensity, which would presumably interact with the consonantal strengthening effects in enhancing the C/V boundary.

In summary, the Accent effects observed here reveal a general pattern of prosodic strengthening for stop consonants. The effects of Accent on the acoustic properties that encode Voice and POA are contrast preserving, in that strengthening effects do not obliterate or reduce phonological contrasts. There is some evidence of contrast enhancement under Accent for the Voice contrast, while acoustic correlates of POA show uniform strengthening with no overall enhancement of the POA contrast.

Contrary to the findings on the effect of Accent, we find no effect of IP-position on the acoustic voicing cues for unaccented /t/ and /d/, or for the combined group of accented and unaccented /t,d/ data. Comparison of standard deviations in IP-Initial and IP-medial /t, d/ for each acoustic measure shows less variability in IP-initial position, suggesting that initial consonants are produced with greater precision in their laryngeal and supralaryngeal articulations. Reduced variability of acoustic cues to phonological contrasts can be viewed as a different kind of strengthening, if the result is greater perceptual salience of those contrasts for the listener.

The absence of a main effect of IP-position on acoustic measures is surprising in light of the evidence for articulatory strengthening in IP-initial position reported in Cho (2001, 2005). Cho (2001) shows evidence of domain-initial strengthening of consonants in #CV structures, corroborating the results of Fougeron and Keating (1997), but as no acoustic data are provided, direct comparison of the present study with his findings is not possible. Cho (2005) provides both articulatory and acoustic evidence of strengthening for the

vowels/a, i/ in domain-initial #CV structures, but notes that “the boundary effect is generally more robust domain-finally than domain-initially” (p.3876). He attributes the weaker effect of domain-initial position on vowels to the fact that target vowels were not in the absolute domain-initial position, and thus may experience a weaker effect than the domain-initial consonant. This explanation does not extend to the present study, however, because the domain-initial effect is examined for the consonants /t,d/, which are in the absolute initial position. Differences between Cho’s findings and ours may be due to differences in the speech materials; his data are from single-sentence, read utterances of the sort that are typical of laboratory studies, while our data are taken from longer narratives read by professional radio news announcers.

5. Conclusion

This study of Radio News speech finds evidence for the strengthening of word-initial stop consonants in accented position in American English. Accent affects the acoustic properties that express the phonological Voice contrast, with the overall result that the Voice contrast is enhanced for stops in accented syllables compared to unaccented syllables. Accent also affects the acoustic cues to phonological POA, and though the effect is variable across POA groups, there is no overall pattern of enhancement of the POA contrast. The only observed enhancement of POA cues was in the increased distinction of labials from non-labials in VOT under Accent. Further, there is no evidence that Accent conditions prosodic enhancement through reduced variability of the acoustic measures that encode POA.

An interesting finding from this study is that accent effects are generally contrast-preserving. Accent effects are greatest for those acoustic properties that play a lesser role in encoding a phonological contrast, suggesting that the phonological function of an acoustic measure outweighs its pragmatic function, at least for the acoustic correlates of phonological Voice and POA contrasts examined here.

Finally, this study did not find evidence of an increased acoustic distinctiveness between phonologically contrastive stops in initial position of an Intonational Phrase compared to medial position. The kind of prosodic strengthening found for Voice and POA cues under Accent is not observed under the phrase-initial boundary condition. On the other hand, acoustic variability appears to be less in initial position suggesting that initial consonants are produced with greater precision in their laryngeal and supralaryngeal articulations, and this reduced variability can be seen as a different kind of prosodic strengthening effect. In the present study, no comparable reduction in variability is found under Accent, which indicates a fundamental difference between the acoustic effects of accent and phrasal position on segment level phonological contrasts.

Acknowledgments

Support for this research came from the University of Illinois through a Critical Research Initiative award to J. Cole and M. Hasegawa-Johnson. The authors acknowledge very helpful comments from Marija Tabain and the editor, and from an anonymous reviewer, and are grateful to Letania Ferreira for assistance with acoustic labeling.

Appendix. A

The frequency of individual vowels (optional) following the studied stop consonants and the frequency of stop tokens in function and nonfunction words for speaker F3A are displayed in [Table A1](#). Descriptive statistics for acoustic measures of stop voicing for stops in IP-medial position, for each speaker is provided in [Table A2](#). Descriptive statistics for VOT, f_0 , CD and BA as measures of stop POA for stops in IP-medial position are displayed for speakers F3A and F2B in [Table A3a](#), and for speakers M1B and M2B in [Table A3b](#). Descriptive statistics for burst spectrum measures in IP-medial position for speakers F3A and F2B and for speakers M1B and M2B are provided in [Tables A4a](#), and [A4b](#), respectively.

Table A1

Frequency of individual vowels (based on lexical values) following the studied stop consonants and frequency of stop tokens in function and nonfunction words. Speaker F3A

		b	d	g	k	p	t
V2	aa	7	2	2	18	5	2
	ae	3	0	1	5	1	4
	ah	10	4	4	2	4	5
	ao	0	0	0	7	0	0
	aw	1	1	0	3	0	0
	ax	0	1	0	15	4	0
	axr	0	0	0	2	1	0
	ay	14	0	0	0	1	5
	eh	1	3	9	0	0	3
	er	6	1	0	1	10	0
	ey	1	1	0	4	0	4
	ih	4	3	2	2	1	0
	ij	10	0	0	0	0	0
	ix	2	11	0	0	0	0
	iy	17	4	0	4	5	2
	ow	1	3	2	0	1	0
	oy	0	0	0	0	2	0
	uh	0	0	0	1	3	0
uw	0	7	0	0	0	65	
function-wd	fwd	45	0	0	4	6	62
	non-fw	32	41	20	60	32	28

Table A2

Descriptive statistics for acoustic measures of stop voicing for stops in IP-medial position, for each speaker individually

		Speaker F3A				Speaker F2B			
Voicing		Mean	s.d.	N	Missing	Mean	s.d.	N	Missing
VOT (ms)	Voiced	19.4	10.3	19	94	17.7	7.7	67	37
	Voiceless	49.4	26.0	168		55.5	21.6	144	
f ₀ (Hz)	Voiced	191.57	27.41	113	0	157.57	29.82	102	8
	Voiceless	213.86	32.43	168		168.46	28.44	138	
CD (ms)	Voiced	65.6	26.1	91	64	55.3	19.7	82	73
	Voiceless	48.9	22.2	126		50.4	16.4	93	
BA (dB)	Voiced	55.23	4.12	111	2	56.96	4.90	102	2
	Voiceless	55.67	6.09	168		56.37	5.57	144	
		Speaker M1B				Speaker M2B			
VOT (ms)	Voiced	16.0	6.4	80	53	21.2	9.5	32	81
	Voiceless	47.1	19.2	175		66.1	31.5	130	
f ₀ (Hz)	Voiced	113.66	13.51	131	8	123.34	31.00	110	10
	Voiceless	120.62	14.98	169		140.89	36.76	123	
CD (ms)	Voiced	59.3	20.4	101	82	66.5	26.3	82	63
	Voiceless	52.1	17.4	125		57.8	21.0	98	
BA (dB)	Voiced	53.04	4.07	125	8	54.65	5.17	102	11
	Voiceless	52.86	3.31	175		48.49	5.61	130	

Table A3a

Descriptive statistics for VOT, f_0 , CD and BA as measures of stop place of articulation for stops in IP-medial position, speakers F3A and F2B

	POA	Speaker F3A				Speaker F2B			
		Mean	s.d.	N	Missing	Mean	s.d.	N	Missing
VOT(ms)	Labial	44.3	25.9	41	94	31.3	21.7	60	37
	Alveolar	37.8	18.4	82		41.0	23.4	82	
	Velar	58.7	31.0	64		57.2	24.8	69	
f_0 (Hz)	Labial	196.96	29.19	100	0	156.88	26.85	89	8
	Alveolar	205.80	30.81	106		165.65	32.88	83	
	Velar	214.20	36.10	75		170.71	26.68	68	
CD (ms)	Labial	66.5	25.8	84	64	61.5	16.9	72	73
	Alveolar	49.5	23.7	70		46.0	18.5	53	
	Velar	48.8	21.3	63		47.1	14.1	50	
BA (dB)	Labial	52.71	4.23	98	2	54.76	4.75	89	2
	Alveolar	59.64	4.69	106		60.52	4.07	87	
	Velar	53.28	3.82	75		54.11	4.52	70	

Table A3b

Descriptive statistics for VOT, f_0 , CD and BA as measures of stop place of articulation for stops in IP-medial position, speakers M1B and M2B

		Speaker M1B				Speaker M2B			
		Mean	s.d.	N	Missing	Mean	s.d.	N	Missing
VOT (ms)	Labial	23.8	013.5	61	53	41.1	027.0	47	81
	Alveolar	36.3	019.5	117		58.9	030.2	62	
	Velar	49.6	023.6	77		69.7	037.3	53	
f_0 (Hz)	Labial	115.58	13.21	104	8	125.20	28.95	90	10
	Alveolar	117.93	15.44	116		131.82	32.18	83	
	Velar	119.67	15.45	80		144.80	43.97	60	
CD (ms)	Labial	69.6	016.6	78	82	75.6	020.7	75	63
	Alveolar	48.3	017.4	86		48.9	018.4	60	
	Velar	47.0	013.3	62		55.9	023.5	45	
BA (dB)	Labial	52.03	4.37	98	8	53.77	6.053	87	11
	Alveolar	53.32	2.87	122		48.61	5.53	83	
	Velar	53.46	3.55	80		51.05	5.92	62	

Table A4a

Descriptive statistics for burst spectrum measures in IP-medial position, speakers F3A and F2B

	POA	Speaker F3A			Speaker F2B		
		Mean	s.d.	N	Mean	s.d.	N
BS–low (dB)	Labial	51.25	4.35	100	53.05	4.56	89
	Alveolar	50.39	5.15	106	50.33	4.63	87
	Velar	47.92	5.23	75	49.64	5.42	70
BS–mid (dB)	Labial	37.99	4.94	100	41.59	5.29	89
	Alveolar	46.15	4.30	106	46.22	4.22	87
	Velar	44.80	4.41	75	44.88	4.72	70
BS–hi (dB)	Labial	32.72	6.20	100	37.80	5.53	89
	Alveolar	53.34	5.10	106	53.73	5.08	87
	Velar	42.37	4.96	75	42.54	6.83	70
BSRatio–low	Labial	.74	.09	100	.79	.10	89
	Alveolar	.92	.12	106	.92	.10	87
	Velar	.94	.13	75	.91	.10	70
BSRatio–hi	Labial	1.18	.13	100	1.11	.10	89
	Alveolar	.87	.10	106	.87	.11	87
	Velar	1.07	.12	75	1.07	.14	70

Table A4b

Descriptive statistics for burst spectrum measures in IP-medial position, speakers M1B and M2B

	POA	Speaker M1B			Speaker M2B		
		Mean	s.d.	N	Mean	s.d.	N
BS–low (dB)	Labial	51.85	6.39	101	54.29	6.74	87
	Alveolar	48.48	4.61	122	49.70	6.07	83
	Velar	50.21	4.52	80	49.98	6.63	62
BS–mid (dB)	Labial	38.20	7.00	101	43.09	6.02	87
	Alveolar	40.79	3.59	122	41.71	3.96	83
	Velar	44.23	3.84	80	40.91	5.77	62
BS–hi (dB)	Labial	31.15	7.04	101	33.65	4.98	87
	Alveolar	43.77	3.88	122	44.71	4.29	83
	Velar	35.58	5.44	80	38.00	5.08	62
BSRatio–low	Labial	.74	.08	101	.80	.08	87
	Alveolar	.85	.09	122	.85	.10	83
	Velar	.89	.11	80	.83	.14	62
BSRatio–hi	Labial	1.24	.12	101	1.30	.22	87
	Alveolar	.94	.11	122	.94	.11	83
	Velar	1.27	.19	80	1.08	.12	62

References

- Beckman, M. E., & Cohen, K. B. (2000). Modeling the articulatory dynamics of two levels of stress contrast. In M. Horne (Ed.), *Prosody: Theory and experiment* (pp. 169–200). Dordrecht: Kluwer.
- 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.

- Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. Docherty, & D. R. Ladd (Eds.), *Papers in laboratory phonology II: gesture, segment, prosody* (pp. 68–86). Cambridge: Cambridge University Press.
- Beckman, M. E., & Elam, G. A. (1997). *Guidelines for ToBI labeling*, version 3 from Ohio State University: http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/singer_tobi.html
- Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255–309.
- Benki, J. R. (2001). Place of articulation and first formant transition pattern both affect perception of voicing in English. *Journal of Phonetics*, 29.1, 1–22.
- Blumstein, S., & Stevens, K. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America*, 66(4), 1001–1017.
- Boersma, P., & Weenink, D. (2006). Praat: Doing phonetics by computer (Version 4.4.23) [Computer program]. Retrieved June 2, 2006, from <http://www.praat.org>
- Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica*, 57(1), 3–16.
- Byrd, D., Kaun, A., & Narayanan, S. (1996). Prosodic boundary effects in Tamil: an articulatory study. Presented at the January, 1996 meeting of the Linguistic Society of America, San Diego, CA.
- Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, 22, 129–159.
- Cho, T. (2001). *Effects of prosody on articulation in English*. PhD dissertation, UCLA.
- Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, 32(2), 141–176.
- Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realization of /a, i/ in English. *Journal of the Acoustical Society of America*, 117(6), 3867–3878.
- Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip movement kinematics in English. In: L. Goldstein, D. Whalen, & C.T. Best (Eds.), *Laboratory phonology 8* (pp. 519–548). Berlin: Mouton de Gruyter.
- Cho, T., & Jun, S-A. (2000). Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistic Society*, 36, 31–44.
- Cho, T., & Keating, P. A. (2001). Articulatory and acoustics studies of domain-initial strengthening in Korean. *Journal of Phonetics*, 29, 155–190.
- Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, 27, 207–229.
- Cho, T., & McQueen, J. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33, 121–157.
- Choi, H. (2003). Prosody induced acoustic variation in English stop consonants. *Proceedings of the 15th international congress of phonetic sciences*, Barcelona, Spain (pp. 2661–2664).
- Cole, J., Choi, H., & Kim, H. (2003). Acoustic evidence for the effect of accent on CV coarticulation in Radio News speech. In A. Agwuele, W. Warren, & S. Park (Eds.), *Proceedings of the 2003 Texas Linguistics conference* (pp. 62–72). Somerville, MA: Cascadilla Proceedings Project.
- Crystal, T. H., & House, A. S. (1982). Segmental durations in connected speech signals: Preliminary results. *Journal of the Acoustical Society of America*, 72, 705–716.
- Crystal, T. H., & House, A. S. (1988a). Segmental durations in connected-speech signals: Current results. *Journal of the Acoustical Society of America*, 83, 1553–1573.
- Crystal, T. H., & House, A. S. (1988b). Segmental durations in connected-speech signals: Syllabic stress. *Journal of the Acoustical Society of America*, 83, 1574–1585.
- de Jong, K. (1991). *The oral articulation of English stress accent*. PhD dissertation, Ohio State University.
- de Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, 97, 491–504.
- Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, 26, 45–69.
- Docherty, G. (1992). *The timing of voicing in British English obstruents*. Berlin, New York: Foris.
- Dorman, M. F., & Raphael, L. J. (1980). Distribution of acoustic cues for stop consonant place of articulation in VCV syllables. *Journal of the Acoustical Society of America*, 67(4), 1333–1335.
- Duez, D. (1992). Second formant locus-nucleus patterns: An investigation of spontaneous French speech. *Speech Communication*, 11, 417–427.
- Edwards, J., Beckman, M., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America*, 89(1), 369–382.
- Erickson, D. (2002). Articulation of extreme formant patterns for emphasized vowels. *Phonetica*, 59, 134–149.
- Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29, 109–136.
- Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, 101, 3728–3739.
- Greenberg, S., Chang, S., & Hitchcock, L. (2001). The relation between stress accent and vocalic identity in spontaneous American English discourse. *Proceedings of ISCA workshop on Prosody in speech recognition and understanding*, Red Bank, NJ. (pp. 51–56).
- Haggard, M., Ambler, S., & Callow, M. (1970). “Pitch as a voicing cue”. *Journal of the Acoustical Society of America*, 47(2.2), 613–617.

- Harrington, J., Fletcher, J., & Beckman, M. (2000). Manner and Place conflicts in the articulation of accent. In M. Broe, & J. Pierrehumbert (Eds.), *Papers in laboratory phonology V: Acquisition and the lexicon* (pp. 40–51). Cambridge: Cambridge University Press.
- Harrington, J., Fletcher, J., & Roberts, C. (1995). Coarticulation and the accented/unaccented distinction: Evidence from jaw movement data. *Journal of Phonetics*, 23, 305–322.
- Hsu, C.-S., & Jun, S.-A. (1998). Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic? *UCLA Working Papers in Phonetics*, 96, 69–89.
- Jessen, M. (1998). *Phonetics and phonology of Tense and lax obstruents in German*. Amsterdam: John Benjamins.
- Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody*. Ph.D. dissertation, Ohio State University.
- Jun, S.-A. (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell, & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Papers in laboratory phonology IV* (pp. 235–253). Cambridge: Cambridge University Press.
- Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, 60, 286–319.
- Keating, P. A., Cho, T., Fougeron, C., & Hsu, C.-S. (2004). Domain-initial articulatory strengthening in four language. In J. Local, R. Ogden, R. Temple, M. E. Beckman, & K. Kingston (Eds.), *Phonetic interpretation: Papers in Laboratory Phonology VI* (pp. 145–163). Cambridge: Cambridge University Press.
- Kewley-Port, D. (1983). Time-varying features as correlates of place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 73, 322–335.
- Klatt, D. H. (1974). The Duration of [s] in English words. *Journal of Speech and Hearing Research*, 17, 51–63.
- Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research*, 18, 686–706.
- Lahiri, A., Gewirth, L., & Blumstein, S. E. (1984). A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study. *Journal of the Acoustical Society of America*, 76, 391–404.
- Lieberman, A. M., Delattre, P. C., & Cooper, F. S. (1958). Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech*, 1, 153–167.
- Lindblom, T. (1990). Explaining phonetic variation: a sketch of the H & H theory. In W. J. Hardcastle, & N. Hewlett (Eds.), *Speech production and speech modeling* (pp. 403–439). Dordrecht: Kluwer.
- Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language*, 33, 42–49.
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 385–422.
- Lisker, L., & Abramson, A. S. (1967). The voicing dimension: Some experiments in comparative phonetics: In *Proceedings of the sixth international congress of phonetic sciences, Prague* (Academia, Prague, 1970), (pp. 563–567).
- Luce, P. A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *Journal of the Acoustical Society of America*, 78, 1949–1957.
- Massaro, D. W., & Oden, G. C. (1980). Evaluation and integration of acoustic features in speech perception. *Journal of the Acoustical Society of America*, 67(3), 996–1013.
- Ohde, R. N., & Stevens, K. N. (1983). Effect of amplitude on the perception of stop consonant place of articulation. *Journal of the Acoustical Society of America*, 74, 706–714.
- Ostendorf, M., Price, P.J., & Shattuck-Hufnagel, S. (1995). *The Boston University Radio News Corpus*, from Linguistics Data Consortium: www ldc.upenn.edu.
- Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty, & D.R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 90–119). Cambridge: Cambridge University Press.
- Plauché, M. C., & Sönmez, K. (2000). Machine learning techniques for the identification of cues for stop place. *Proceedings of ICSLP 2000, Beijing, China* (pp. 548–551).
- Port, R. (1976). The influence of speaking tempo on the duration of stressed vowel and medial stop in English trochee words. Unpublished doctoral dissertation, University of Connecticut.
- Port, R. F. (1979). The influence of tempo on stop closure duration as a cue for voicing and place. *Journal of Phonetics*, 7, 45–56.
- Repp, B. H. (1979). Relative amplitude of aspiration noise as a cue for syllable-initial stop consonants. *Language and Speech* 22, 173–179.
- Repp, B. H., & Lin, H. B. (1989). Acoustic properties and perception of stop consonant release transients. *Journal of the Acoustical Society of America*, 85(1), 379–396.
- Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 64, 1358–1368.
- Stevens, K. N., & Klatt, D. H. (1974). Role of formant transitions in the voiced–voiceless distinction for stops. *Journal of the Acoustical Society of America*, 55, 653–659.
- Summerfield, Q., & Haggard, M. (1977). On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *Journal of the Acoustical Society of America*, 62, 435–448.
- Tabain, M. (2003). Effects of prosodic boundary on /aC/ sequences: acoustic results. *Journal of the Acoustical Society of America*, 113, 516–531.
- Tartter, V. C., Kat, D., Samuel, A. G., & Repp, B. H. (1983). Perception of intervocalic stop consonants: The contributions of closure duration and formant transitions. *Journal of the Acoustical Society of America*, 74(3), 715–725.
- Turk, A. E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics*, 27, 171–206.
- Umeda, N. (1977). Consonant duration in American English. *Journal of the Acoustical Society of America*, 61, 846–858.
- Whalen, D. H., Abramson, A. S., Lisker, L., & Moody, M. (1993). F0 gives voicing information even with unambiguous voice onset times. *Journal of the Acoustical Society of America*, 93, 2152–2159.

- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, *91*, 1707–1717.
- Winitz, H., Scheib, M. E., & Reeds, J. A. (1972). Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech. *Journal of the Acoustical Society of America*, *51*(4.2), 1309–1317.
- Wouters, J., & Macon, M. W. (2002). Effects of prosodic factors on spectral dynamics. I. Analysis. *Journal of the Acoustical Society of America*, *111*, 417–427.
- Zue, V. W. (1976). Acoustic characteristics of stop consonants: A controlled study. Sc.D. Thesis, MIT.