Variation in stop consonant voicing in two regional varieties of American English

Ewa Jacewicz
Robert Allen Fox
Samantha Lyle

Speech Perception and Acoustics Laboratories,
Department of Speech and Hearing Science,
The Ohio State University
jacewicz.1@osu.edu, fox.2@osu.edu, samantha.lyle@gmail.com

This study is an acoustic investigation of the nature and extent of consonant voicing of the stop /b/ in two dialectal varieties of American English spoken in south-central Wisconsin and western North Carolina. The stop /b/ occurred at the juncture of two words such as *small bids*, in a position between two voiced sonorants, i.e. the liquid /l/ and a vowel. Twenty women participated, ten representing the Wisconsin and ten the North Carolina variety, respectively. Significant dialectal differences were found in the voicing patterns. The Wisconsin stop closures were usually not fully voiced and terminated in a complete silence followed by a closure release whereas North Carolina speakers produced mostly fully voiced closures. Further dialectal differences included the proportion of closure voicing as a function of word emphasis. For Wisconsin speakers, the proportion of closure voicing was smallest when the word was emphasized and it was greatest in non-emphatic positions. For North Carolina speakers, the degree of word emphasis did not have an effect on the proportion of closure voicing. The results suggest different mechanisms by which closure voicing is maintained in these two dialects, pointing to active articulatory maneuvers in North Carolina speakers and passive in Wisconsin speakers.

1 Introduction

It has been established in the literature that conditions for voicing in voiced stop consonants in American English are more favorable in the intervocalic position than in word-initial or word-final positions. The intervocalic position enhances the ‘articulatory ease’ (Westbury & Keating 1986) which can facilitate continuous voicing through the oral closure of a single voiced stop (e.g. Westbury 1983). In intervocalic positions, speakers sustain the vibration of the vocal folds longer so that the low-frequency periodic energy lasts for an extended period of time of stop closure duration. In such positions, voicing can even persist into the closure of a voiceless stop up to 30% of closure duration (e.g. Flege & Brown 1982, Docherty 1992).

This paper investigates the nature and extent of voicing during the closure of the voiced stop consonant /b/ in American English when occurring in a position between two voiced sonorants, the preceding liquid /l/ and a vowel that follows the stop. A voiced stop in this intersonorant position shares similar characteristics with the voiced stop in the intervocalic position. In particular, glottal vibration can still continue during the stop closure especially if the volume of the vocal tract increases and the stop closure duration is short. Increases in vocal tract volume can be achieved by such supplementary adjustments as lowering the

Journal of the International Phonetic Association (2009) 39/3
doi:10.1017/S0025100309990156

larynx, advancing the tongue root, depressing the tongue body, or expanding the oral cavity (through enlargement of the cheeks or expansion of lateral pharyngeal walls) and venting the airstream through the nasal cavity (Kent & Moll 1969, Bell-Berti 1975, Ohala & Riordan 1979, Svirsky et al. 1997). During the closure, both the frequency of glottal vibration and the amplitude of the glottal pulses decrease because of the reduced transglottal pressure (Stevens 1998).

Acoustically, stop closure voicing has been studied primarily as a cue which distinguishes between voiced and voiceless stops in languages such as French, German, and English, particularly when the stop occurs in word-final (or syllable-final) position (e.g. Snoeren, Hallé & Segui 2006). However, the other cues associated with the voicing distinction in stops and not the closure voicing per se became the primary research focus. For example, it has been demonstrated that syllables with voiced final stops have longer vowels than those with voiceless final stops (e.g. Chen 1970, Crowther & Mann 1992). For stops in word-initial (or syllable-initial) positions, it has been shown that phonologically voiced stops in American English are produced most of the time without voicing portion during the closure. Rather, voicing begins shortly after the closure release (e.g. Lisker & Abramson 1964) and other acoustic cues such as higher and lower fundamental frequency (F0) values contribute to the distinction between voiceless and voiced stops, respectively (Ohde 1984). In intervocalic positions, voiced stops have been found to have shorter closures, a longer presence of vocal folds vibration during closure, and lower F0 and F1 at the edges of neighboring vowels as compared to their voiceless counterparts (e.g. Lisker 1957, 1986). The lower F0 values in adjacent vowels come from a deliberate slackening of the vocal folds whereas the lower frequency of the onset of F1 is a function of the timing of voice onset (see Kingston & Diehl 1994).

This brief sketch underscores the primary focus of most studies which inquired into the stop voicing: To reveal how the voicing CONTRAST in stops is phonetically cued to maintain the phonological [+/– voice] distinction. However, little work has been done to examine possible systematic differences in the production of voicing patterns WITHIN each ‘voicing category’ of the stop. For example, considering American English voiced stops only, there is evidence from developmental studies that young children’s voiced stops are often devoiced and have less closure voicing than adults (Kewley-Port & Preston 1974, Smith 1979). Recently, Koenig & Lucero (2008) found notable variation in the proportion of closure voicing in American English-speaking children and female speakers. For the voiced stop /b/ in intervocalic position, the proportion of closure voicing increased over age so that it was 36% for the 5-year-olds, 59% for the 10-year-olds, and 64% for women. No such systematic effect, however, was found for the voiceless stop /p/. Developmental factors seemed to target the voiced stops in the production of the youngest children only, whose /b/ was devoiced most often and the stop had least closure voicing as compared to older children and adults.

Given the possibility of variable production of voiced stops in intervocalic position for the speakers of the same language, the question arises whether there can be systematic differences in closure voicing related to different dialectal backgrounds of the speakers. Although studying the cross-dialectal variation in the realization of the voicing contrast (i.e. comparing the phonetic realization of both voicing categories) will be of primary focus of cross-dialectal research to come, the present socio-phonetic investigation focuses somewhat narrowly on the closure voicing of a single voiced stop. The set of most effective measures developed in this study will serve as a basis for further investigation of voicing patterns addressing a broader range of linguistic inquiries.

There is an abundance of evidence that regionally defined variation in American English introduces considerable differences in pronunciation patterns across the country. However, the majority of acoustic studies of segmental variation in American English dialects have examined vowels and diphthongs (e.g. Labov, Jaeger & Steiner 1972; Thomas 2000, 2001; Clopper, Pisoni & de Jong 2005; Jacewicz, Fox & Salmons 2006; Labov, Ash & Boberg 2006). In Thomas’ (2002: 171) words, ‘consonants, prosody, and voice quality remain, for the most part, in the realm of impressionistic phonetics’. Among the sparse acoustic investigations into stop consonant voicing, Purnell, Salmons & Tepeli (2005) and Purnell, Salmons, Tepeli & Mercer (2005) found considerable variation in the production of stops in the word-final position in Wisconsin. Apparently, in parts of Wisconsin, speakers produce the voicing distinction in the final position differently from what has been reported for American English, showing a partial neutralization of the contrast. This pattern has most likely its roots in the German heritage of the Wisconsin speakers.

The present study is an acoustic investigation of consonant voicing in two very different regional varieties of American English found in western North Carolina (representing the southern American English often referred to as Appalachian English) and in central Wisconsin (representing the northern variety). Although the vowel systems of these two dialects are widely known to undergo two distinct chain shifts, namely the Southern Vowel Shift in North Carolina and Northern Cities Shift in Wisconsin (Labov 1994), we know considerably less about possible differences in the pronunciation of their consonants. Yet, impressionistic evaluation of speech samples in our labs produced by speakers from the South and from the North has led to the observation that many bilabial stops /b/ in southern speech are strongly pre-voiced and sound very sonorous, sometimes even slightly pre-nasalized. There was also some evidence of an 'implosive' characteristic to their release. In contrast, the same stops in the production of Wisconsin speakers appeared weak and almost devoiced. Most importantly, these differences have been observed for the intersonorant position of the stop, which is not affected by additional factors stemming from word-final or word-initial positions.

If such dialectal differences do exist, the breath-stream control mechanism cited earlier (Westbury & Keating 1986) which was proposed on the basis of American English may not be as effective in the production of North Carolina stops. Specifically, according to this model, a stop occurring in the intervocalic position should naturally be voiced throughout most if not all of the closure if it is short. In the case of a longer closure, the stop will be voiced through the initial portion of the closure and then become voiceless. Does this prediction hold for all dialects of American English? The present acoustic investigation is directed toward a better understanding of these voicing patterns. A direct comparison of the southern and northern productions of an American English stop in the intersonorant position will provide first cross-dialectal data which may initiate further interest in testing the existing models of speech production.

2 Methods

2.1 Speakers

Twenty women aged between 51 and 65 years participated in the study (recordings were made in 2006 and 2007). There were 10 speakers from each dialect area who were born, and raised, and spent most of their lives, in either south-central Wisconsin (the Madison area: Dodge and Dane counties) or western North Carolina (the Sylva, Cullowhee, and Waynesville areas: Jackson, Swain and Haywood counties). The Wisconsin speakers reported growing up in a suburban setting and the North Carolina speakers grew up mostly in rural areas or small towns. It needs to be pointed out that these small towns in North Carolina are located close to the Great Smoky Mountains National Park, a popular tourist destination and are also in close proximity to the Western Carolina University in Cullowhee. The present Wisconsin and North Carolina speakers do not reflect a clear urban/rural split. From a review of their background questionnaires regarding demographics, education and socio-economic status, the subject populations in this study are generally similar in basic demographic terms. This also includes their education levels (mostly some college and high school). None of the participants was bilingual or had any extensive experience using a foreign language. None reported a speech disorder.

2.2 Stimuli

Stimulus material consisted of sets of sentence pairs which contained the target words of the structure /b**V**ts/ and /b**V**dz/, where V was one of the following vowels: /ɪ ɛ æ e aɪ/. The target words were: *bits*/*bids*, *bets*/*beds*, *bats*/*bads*, *baits*/*bades*, *bites*/*bides*. Their position within a sentence did not change nor did the immediate phonetic context in which the target words occurred. However, several different emphasis conditions for each target word were created to introduce variability coming from suprasegmental structure, which is different from variability as a function of immediate phonetic context. To do so, main sentence stress was systematically varied for each sentence set. In this way, the proximity of the target word to the main sentence stress determined the level of emphasis of the target word such as high, intermediate, or low (see Jacewicz et al. 2006 for additional discussion). As illustrated below, for each target word in bold, the variable main sentence stress (marked for the speaker by way of capitalization to elicit accurate production) yielded three gradient levels of target word emphasis:

HIGH
Ted thinks the fall SALES are low.
No! Ted thinks the fall **BIDS** are low.

INTERMEDIATE
Ted thinks the SPRING bids are low.
No! Ted thinks the FALL **bids** are low.

LOW
Ted thinks the fall bids are HIGH.
No! Ted thinks the fall **bids** are LOW.

To determine these three levels of emphasis, more sentence pairs were initially elicited from each speaker, in which the main sentence stress varied systematically throughout the seven-syllable phrase such as:

TED thinks the fall bids are low.
Ted THINKS the fall bids are low.
Ted thinks the FALL bids are low.
Ted thinks the fall BIDS are low.
Ted things the fall bids are LOW.

The sentences with inaccurate stress placement were rejected in the process of data collection and re-recorded to ensure that each capitalized word was produced as intended. The selection of the three levels of emphasis (High, Intermediate, Low) was then determined on the basis of separate acoustic analyses, including vowel duration and the dynamic formant movement (i.e. formant trajectory length).

The word-initial stop /b/ to be analyzed in this study occurred at the juncture of two words such as *small bids*, in a position between two voiced sonorants, i.e. the liquid /l/ and a vowel. Only the second sentence in each set was analyzed so that the stop /b/ always occurred in the position between the two voiced sonorants. The complete stimulus material is listed in the Appendix. In this study, no acoustic measurements of the preceding liquid /l/ were made because there were no discernable differences in the production of /l/ for the speakers of the two dialects. It might be the case that the preceding liquid /l/ could potentially influence vocal folds vibration during the stop closure due to the variation in the degree of velarization of /l/. We controlled for this factor by selecting words containing only open back vowels before /l/, i.e. *small*, *dull*, *fall*, and *tall*, which to some extent reduced the vowel–liquid transitional effects on the quality of the /l/-sound. In addition, only speakers in the same age range were selected for the present analysis. Although we were unaware of any potential changes in the pronunciation of /l/ across the lifespan at the time of data collection, it has been shown recently that the degree of velarization can change across generations of speakers (Van Hofwegen & Thomas 2009).

2.3 Recording procedure

Recordings were made under computer control using a program written in Matlab. Sentence pairs were presented in random order on a computer screen and read by a speaker seated in a sound-attenuating booth. The stimuli were recorded directly onto a hard drive disk at a 44.1 kHz sampling rate. A head-mounted microphone (Shure SM10A) was used, placed one inch from the lips. Each speaker was instructed to read the sentences fluently and place the main sentence stress on the capitalized words. The recordings were repeated in case of mispronunciations, wrong placement of the main sentence stress, or disfluent productions. Two research assistants helped with data collection, one in Wisconsin and one in North Carolina so that all participants in a given state were recorded by the same experimenter. A total of 60 sentences were obtained from each speaker (5 vowels × 2 final consonantal contexts × 3 emphatic positions × 2 repetitions). In total, 1200 word tokens were analyzed from all twenty speakers.

2.4 Acoustic measurements

Prior to analysis, the tokens were downsampled to 11.025 kHz and pre-emphasized (98%). As already mentioned, only the target word in the second sentence in each sentence pair was analyzed in the current study. The locations of several acoustic landmarks were measured for both the consonant /b/ and the target word itself. These landmarks included the onset of the stop closure, closure release, voicing offset during stop closure, voicing onset for the vowel, word onset (which was the same as the stop closure onset) and word offset. Measurements were made by hand from the waveform displayed in Adobe Audition 1.0 analysis and editing program on one computer screen and verified with spectrograms in the TF32 analysis program (Milenkovic 2003) displayed on the second screen. A custom Matlab program was then used which read all input values and showed the segmentation marks superimposed over a display of the word's waveform in order to make corrections as needed. This program, while performing a measurement check, calculated all output values for further statistical analysis. A second reliability check of all landmark locations using the same Matlab program was performed on all tokens by an experienced researcher.

Stop closure onset was located at a zero-crossing after acoustic energy of the preceding /l/ was greatly reduced. Its location was based on three cues: a drop in overall amplitude of the waveform, a drop in intensity of the formants and a change in the waveform period to become more sinusoidal (due to reduction in energy of higher formants). Voicing offset during the consonantal closure and release were identified visually from the waveform. Vowel onset was measured from onset of periodicity (at a zero crossing) following the release burst. In those cases where closure remained voiced throughout and there was no visual spectrographic (or waveform) evidence of a release burst, vowel onset was located at the point where amplitude increased significantly and spectrograms showed prominent higher formants. Vowel offset for words ending in a voiceless /ts/ was located at the point at which the amplitude of the vowel dropped to near zero (which was also coincident with elimination of all periodicity in the waveform). The vowel offset for words ending in a voiced /dz/ was defined as that point when the amplitude dropped to near zero and any periodicity in the waveform contained no high frequency components. Word offset was located at the cessation of noise of the word-final fricative /s/ or /z/.

The acoustic landmark locations of closure onset, voicing offset, closure release, vowel onset and word offset permitted calculations of several measures of interest to this study which included closure duration, closure voicing duration, proportion of the closure that was voiced, as well as word duration and proportion of closure-to-word duration and VOT. All calculations were done using a Matlab program.

2.5 Statistical analysis

Repeated-measures analyses of variance (ANOVAs) were conducted on word duration, stop closure duration, proportion of closure-to-word duration, closure voicing duration, percentage of closure voicing and frequency of voiced-through closure. The within-subject factors were final consonant in the word (referred to throughout the manuscript as /t/ and /d/ for abbreviation purposes although the actual factor was the status of voicing of the entire final consonant cluster /ts/ or /dz/), word emphasis (High, Intermediate, Low) and vowel (/ɪ ɛ æ e aɪ/). Dialect was the between-subject factor. The α level for determining statistical significance was .05. For all reported significant main effects and interactions, the degrees of freedom for the F-tests were Greenhouse-Geisser adjusted in those cases in which there were significant violations of sphericity. In addition to the significance values, a measure of the effect size – partial eta squared (η^2) – is also reported. The value of η^2 can range from 0.0 to 1.0 and it should be considered a measure of the proportion of variance explained by a dependent variable when controlling for other factors.

In the results to be presented below, all statistical tests pertaining to proportion measures (i.e. proportion of closure-to-word duration, proportion of voicing in closure, and proportion of the voiced-through closure in the present sample) were carried out on arcsine transformed RAU values (Studebaker 1985) even though the mean values are reported and discussed as untransformed proportional data. As is well known, the mean of percentage (and proportional) scores are not independent of the variance of these scores (due to ceiling and basement effects). These scores thus violate the homogeneity of variance assumption required in many statistical tests, including the ANOVAs used here. Thus, prior to the analysis of these data, the percentage scores were arcsine transformed into rationalized arcsine units or RAUs. The advantage of Studebaker's method over other arcsine transformations is that the RAUs have values that are numerically close to the original untransformed percentages.

3 Results

Before presenting the results for each measure, it may be useful to examine a few examples that illustrate the commonly encountered differences in closure voicing between a typical Wisconsin (WI) and a typical North Carolina (NC) speaker.

Figure 1 shows waveforms of closures of the stop /b/ in the words *bades* and *baits* produced by a 55-year-old female WI speaker and a 59-year-old female NC speaker. Both speakers read the set of sentences for this study with comparable fluency (i.e. there were no noticeable pauses in their productions) and at a comparable articulation rate, which was 3.23 syllables per second (syll/s) for the WI speaker and 3.17 syll/s for the NC speaker. The waveform displays include examples of stop closures for each degree of word emphasis examined here. The displays are time-aligned so that each waveform begins with a 15 ms final portion of /l/ preceding the stop closure. The closure terminates with a second 15 ms interval measured from the release, which consists of release burst (if present) and a portion of vowel onset.

As can be seen, there are clear differences between the closures of the WI and NC speakers in terms of both voicing and duration. The NC closures are fully voiced whereas the WI closures begin with a period of voicing which ceases gradually and the closure terminates in a complete silence. There is an energetic release burst for this particular WI speaker whereas no such release can be detected in the production of the NC speaker. The longest closure was found in the high emphasis position of the word, followed by intermediate and low, respectively, although the difference between the latter two is rather small. The WI closures tend to be longer than NC closures across all levels of word emphasis.

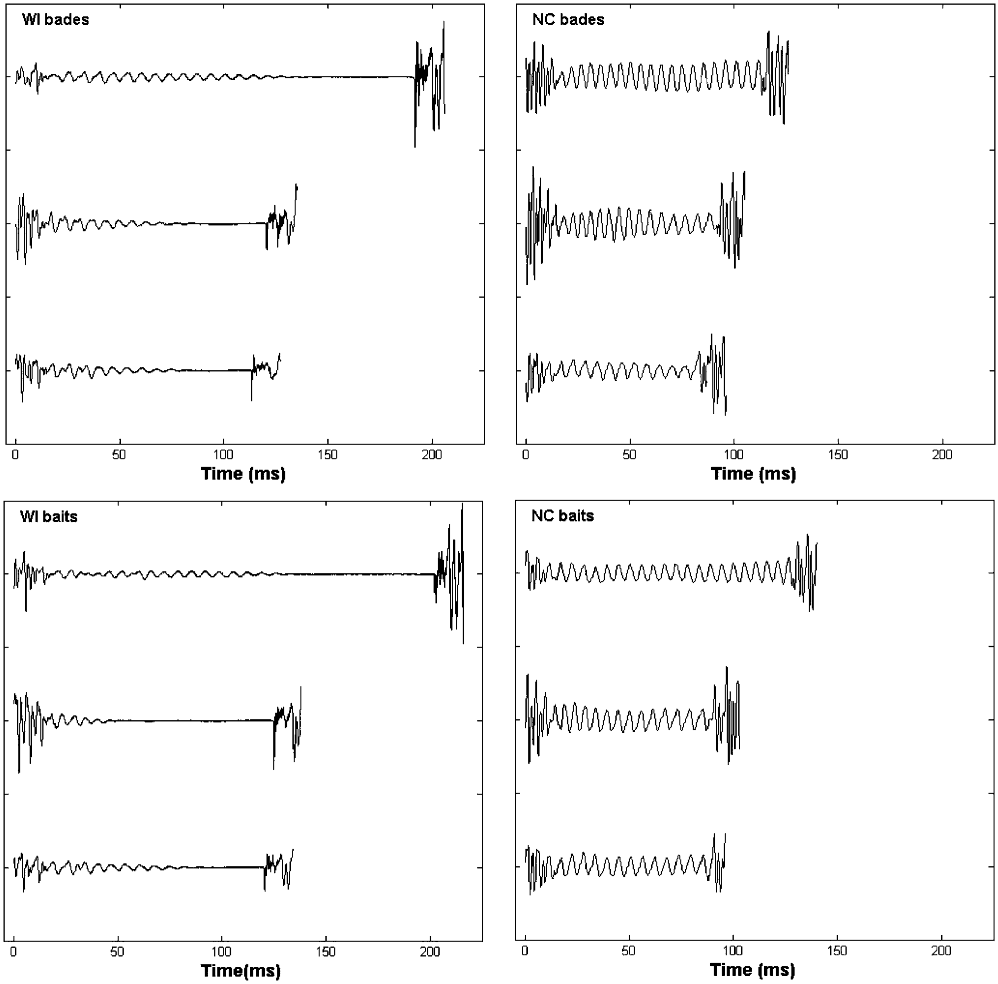

Figure 1 Exemplar waveforms of closures of the stop /b/ in *bades* and *baits* in intersonorant position produced by a female Wisconsin speaker (left) and a female North Carolina speaker (right). In each panel, the stop closures are displayed for each of the three word emphasis positions (counting from the top): High, Intermediate, Low. See text for further details.

With these differences in mind, the results are now presented for each measure selected in this study to assess the general trend and significance of variation in the stop closures produced by WI and NC speakers. Before we address the closure voicing pattern for each dialect, we examine the durations of both the word and the closure as well as the proportion of closure-to-word duration. Next, the nature of voicing during the stop closure is investigated. In particular, we will examine the duration of voicing during the closure, the proportion of closure which is voiced and the frequency with which speakers produced fully voiced stops (i.e. where voicing occurred across the entire duration of the closure) in the present data set.

3.1 Word duration

The overall mean word duration was 422 ms for WI speakers and 464 ms for NC speakers. On average, words produced by NC speakers were 9% longer. However, ANOVA results showed

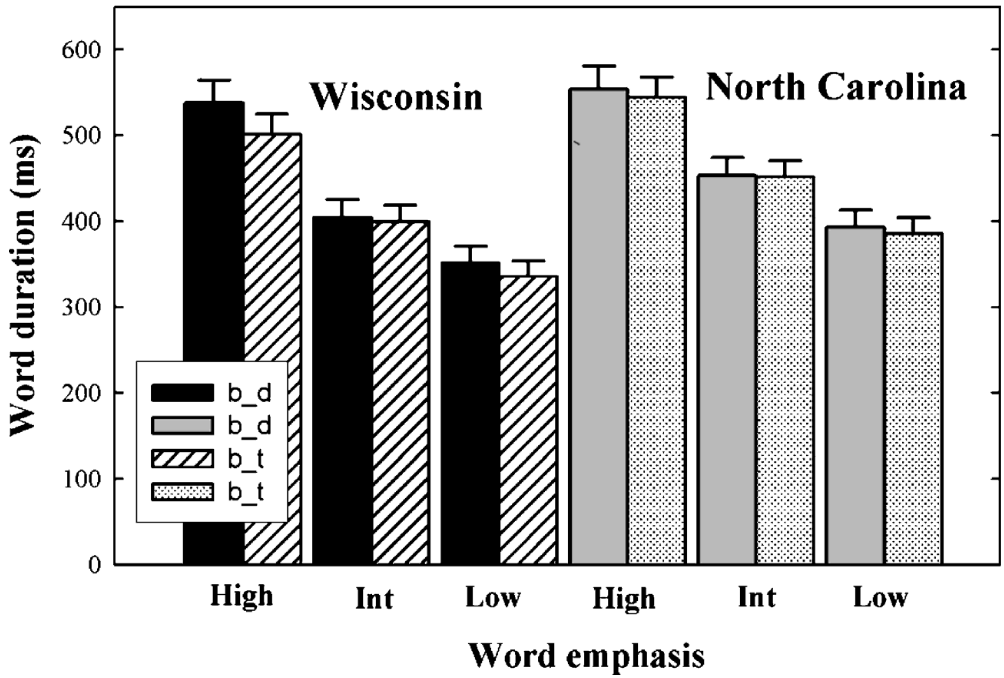

Figure 2 Mean durations of *b_d* and *b_t* words for Wisconsin and North Carolina speakers as a function of word emphasis (High, Intermediate, Low). Here and in all remaining figures, error bars indicate one standard error.

that this trend was not significant. All three within-subject factors were significant. There was a significant main effect of final consonant. Words in the *b_d* context were significantly longer than in the *b_t* context ($F(1,18) = 10.25$, $p = .005$, $\eta^2 = .363$), and the difference can be attributed to the expected vowel duration differences before a voiced and voiceless consonant in English. There was also a strong significant effect of word emphasis ($F(1.6,28) = 101.26$, $p < .001$, $\eta^2 = .849$). On average, words produced with high emphasis were longest (534 ms), followed by intermediate (427 ms) and low emphasis (367 ms), respectively. As shown in figure 2, the effects of word emphasis on word duration are strong for both dialects and persist in both consonantal contexts. There were no significant interaction effects.

There was also a strong significant main effect of vowel ($F(3.5,63.3) = 50.33$, $p < .001$, $\eta^2 = .737$). Words containing one of the short vowels /ɪ ɛ/ were on average shorter (412 ms and 420 ms, respectively) than words containing longer or diphthongal vowels /e æ aɪ/ (457 ms, 461 ms and 463 ms, respectively).

3.2 Stop closure duration

On average, stop closures were longer for WI than for NC speakers (110 ms vs. 101 ms) although this difference did not reach statistical significance. Considering the other factors, there was a significant main effect of word emphasis ($F(1.9,33.3) = 40.26$, $p < .001$, $\eta^2 = .691$). Mean closure duration was longest when the word was produced in high emphasis position and systematically decreased with each lower degree of emphasis. The mean duration values, in descending order of degrees of emphasis, were 137, 99 and 80 ms. Separate ANOVAs making pairwise comparisons between the emphasis positions were used as post-hoc analyses and indicated that all three levels differed significantly from one another. The main effect of vowel was also significant ($F(3,54.1) = 6.18$, $p = .001$, $\eta^2 = .256$). Interestingly, there was

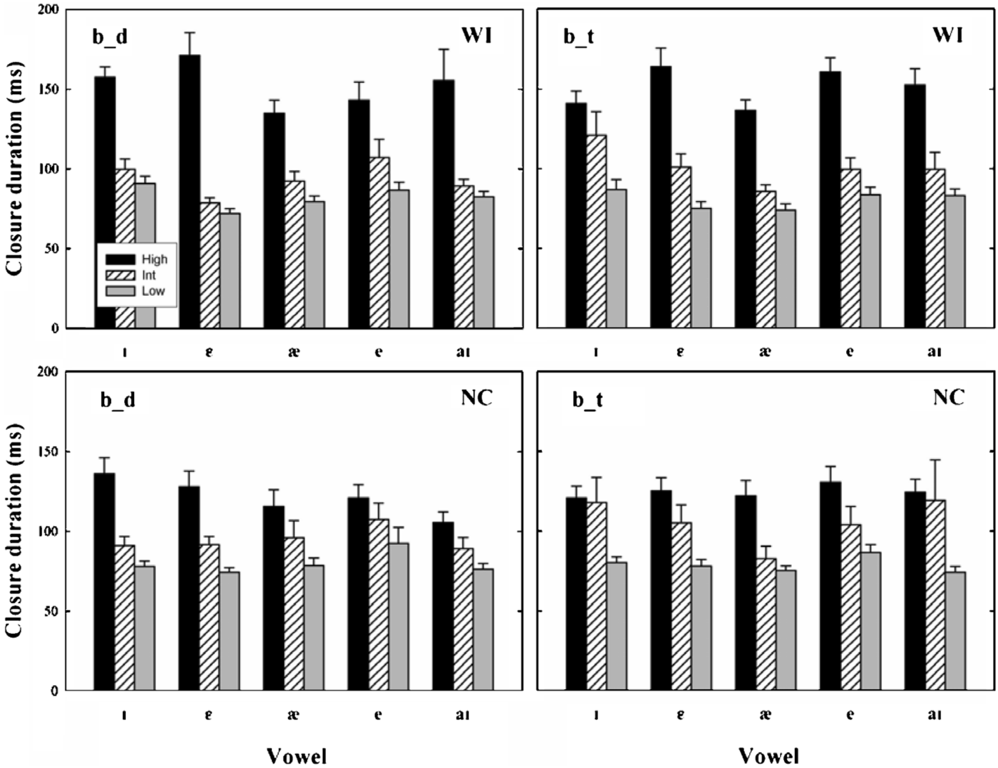

Figure 3 Mean closure durations in *b_d* and *b_t* words shown for each vowel quality and word emphasis position (High, Intermediate, Low) produced by Wisconsin and North Carolina speakers.

no clear relation between the length of the closure and duration of the vowel. The longest closures (m = 110 ms) were found for the short vowel /ɛ/ and the shortest (m = 98 ms) for the long vowel /æ/. There was also a significant three-way interaction between final consonant, stressed position and vowel ($F(3.6,64.6) = 3.55, p = .014, \eta^2 = .165$), which is illustrated for each dialect separately in figure 3. As figure 3 shows, the relation between closure duration, vowel category, word emphasis, and final consonant in the word is relatively complex. The closure duration varies greatly as a function of all these factors although it is noteworthy that, for each dialect, the degree of emphasis affects the closure duration in a systematic way across all vowel categories and word types examined here.

3.3 Proportion of closure-to-word duration ($Prop_{CWD}$)

The opposite trends for word and closure duration, although not significant, suggest that a different approach in comparing the dialects may be more instructive. Measured in absolute terms, closure duration was generally longer for WI speakers although the duration of their words was slightly shorter. This relation may become important if we examine what proportion of the word duration constitutes the stop closure duration. That is, the dialectal differences may exist when closure duration is expressed in relative terms, i.e. as a proportion to the duration of the word (with values ranging from 0% to 100%). The results show that this was indeed the case. The mean $Prop_{CWD}$ was greater for WI speakers (26%) than for NC speakers (22%) and this effect was significant ($F(1,18) = 6.5, p = .020, \eta^2 = .264$). Thus, the longer closure duration and the shorter word duration for WI speakers resulted in a larger $Prop_{CWD}$ for WI as compared to NC.

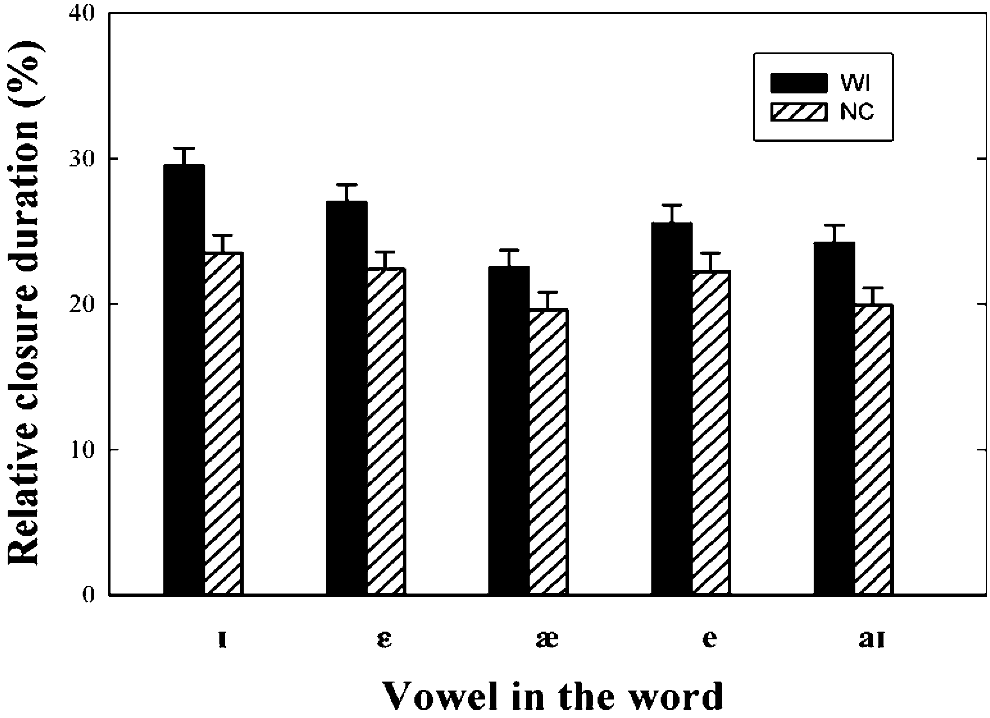

Figure 4 Proportion of closure to word duration shown for each vowel quality produced by Wisconsin and North Carolina speakers.

The $Prop_{CWD}$ was significantly greater in *b_t* words (24%) than in *b_d* words (23%) ($F(1,18) = 34.7$, $p < .001$, $\eta^2 = .658$). This effect was not found for the absolute measure of closure duration. The $Prop_{CWD}$ also varied significantly as a function of word emphasis ($F(1.4,25.5) = 7.3$, $p = .006$, $\eta^2 = .289$). The mean $Prop_{CWD}$ was greatest under high emphasis (26%) followed by intermediate and low, respectively (23% vs. 22%). There was a strong significant effect of vowel ($F(3.8,69.2) = 59.03$, $p < .001$, $\eta^2 = .766$). The $Prop_{CWD}$ was greater when the words contained short vowels /ɪ ɛ/ (27% and 25%, respectively), and smaller when it contained long vowels /e aɪ æ/ (24%, 22% and 21%, respectively). A significant vowel by dialect interaction ($F(3.8,69.2) = 3.74$, $p = .009$, $\eta^2 = .172$) was also obtained. This interaction arose because notable differences to this general pattern were obtained as a function of the dialect, which is illustrated in figure 4. As can be seen, the $Prop_{CWD}$ for words containing NC variants of /æ/ and /aɪ/ are equally low. Similarly, there is no difference in the $Prop_{CWD}$ for words containing the NC vowels /ɛ/ and /e/.

3.4 Closure voicing duration

Against this pattern of variation in stop closure duration, the results for closure voicing are presented below. The duration of the voicing portion during the closure was examined first. As it turned out, mean closure voicing duration for WI speakers was shorter than for NC speakers (69 ms vs. 89 ms) and the main effect of dialect was significant ($F(1,18) = 6.43$, $p = .021$, $\eta^2 = .263$). Significant also was the effect of final consonant in the word ($F(1,18) = 9.34$, $p = .007$, $\eta^2 = .342$). The voicing portion of the closure was shorter in *b_t* words as compared to *b_d* words. However, there was also a significant interaction between the final consonant and dialect ($F(1, 18) = 8.65$, $p = .009$, $\eta^2 = .324$). Closure voicing was significantly shorter

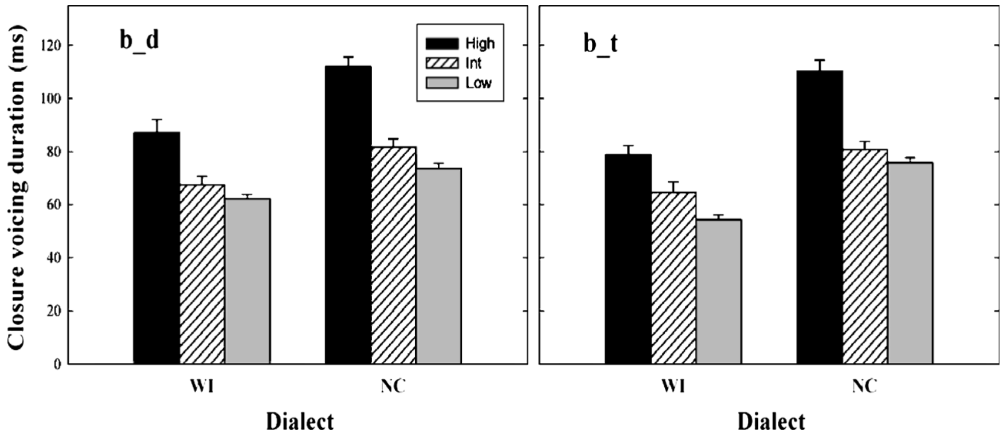

Figure 5 Mean closure voicing duration in *b_d* and *b_t* words for each of the three positions of word emphasis produced by Wisconsin and North Carolina speakers.

in *b_t* than in *b_d* words for WI speakers (66 ms vs. 72 ms) but not for NC speakers (both 89 ms).

Closure voicing duration also varied significantly as a function of word emphasis ($F(1.8,31.8) = 27, p < .001, \eta^2 = .600$). The extent of voicing was greatest in high emphasis position, followed by intermediate and low, respectively. Figure 5 illustrates the effects of word emphasis on closure voicing duration for each dialect. These effects were further explored in separate ANOVAs with the same factors making pairwise comparisons between the emphasis positions. Closure voicing was significantly longer in high emphasis position as compared to low emphasis ($F(1,18) = 46.6, p < .001, \eta^2 = .721$) and it was significantly longer in high emphasis when compared to intermediate emphasis position ($F(1,18) = 22.5, p < .001, \eta^2 = .555$). However, the intermediate and low emphasis positions did not differ significantly from one another ($F(1,18) = 4.1, p = .057, \eta^2 = .187$). This effect was true for each dialect and there was no significant interaction between dialect and emphasis in any of the ANOVAs.

3.5 Proportion of voicing in the closure ($Prop_{VC}$)

Given the differences in closure duration for WI and NC speakers as well as significant dialectal differences in the $Prop_{CWD}$, a cross-dialectal comparison of closure voicing is more appropriate when it is assessed in relative terms, i.e. as the duration of closure voicing relative to the overall duration of the closure itself. An analysis of the proportion of voicing in the closure showed a significant effect of dialect ($F(1,18) = 18.1, p < .001, \eta^2 = .501$). $Prop_{VC}$ was on average smaller for WI speakers than for NC speakers (67% vs. 92%). The main effects of both final consonant and word emphasis were significant ($F(1,18) = 34.1, p < .001, \eta^2 = .655$ and $F(1.5,26.5) = 9.6, p = .002, \eta^2 = .349$, respectively). However, it was the significant interactions between final consonant and dialect and between word emphasis and dialect that shed more light on the dialect-specific changes in the proportion of closure voicing than the main effects themselves.

In particular, the final consonant by dialect interaction ($F(1,18) = 14.2, p = .001, \eta^2 = .441$) showed that the proportion of closure voicing was significantly greater in *b_d* words than in *b_t* words for WI speakers (72% vs. 65%) but not for NC speakers (92% vs. 91%). The second interaction, that of word emphasis by dialect ($F(1.5,26.5) = 5.8, p = .014, \eta^2 = .243$) is illustrated in figure 6. As can be seen, there are clear dialectal differences in the proportion of closure voicing. For WI speakers, there is a relationship between the length of the closure and the $Prop_{VC}$: The shorter the closure duration, the greater $Prop_{VC}$. Thus,

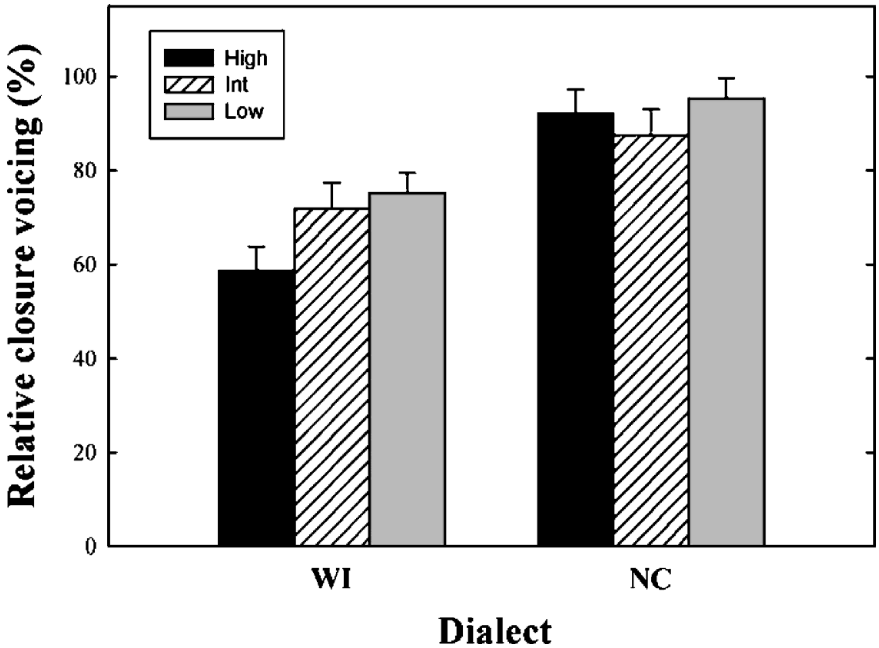

Figure 6 Proportion of voicing in the closure for each of the three positions of word emphasis produced by Wisconsin and North Carolina speakers.

$Prop_{VC}$ is smallest in words in high emphasis position, followed by intermediate and low, respectively. For NC speakers, there is no such relationship and the closure is almost entirely voiced regardless of the variation in closure duration.

3.6 Frequency of fully voiced stop closures

We also examined the frequency with which the stop closures were fully voiced when considering all productions of NC and WI speakers. The results show a strong effect of dialect ($F(1,18) = 24.8$, $p < .001$, $\eta^2 = .580$). The frequency of fully voiced closures was significantly greater for NC speakers (73%) than for WI speakers (24%). There was also a significant main effect of final consonant ($F(1,18) = 8.9$, $p = .008$, $\eta^2 = .331$), showing that fully voiced closures occurred more often in *b_d* words than in *b_t* words (52% vs. 45%). Finally, there was significant effect of word emphasis ($F(1.9,33.9) = 6.84$, $p = .004$, $\eta^2 = .275$). The frequency of fully voiced closures was greatest in the low emphasis position (59%) and decreased in intermediate and high positions, respectively (46% and 41%).

These results clearly show that speakers produce a fully voiced closure more often when the closure is much shorter such as in words that are least emphasized. Also, the anticipation of a final voiced consonant cluster may contribute to 'an ease of articulation' and extensive voicing of the stop.

4 General discussion

In this study of the voicing patterns in voiced stops in two American English dialects, we examined the production of the stop /b/ in intersonorant position. A systematic variation in

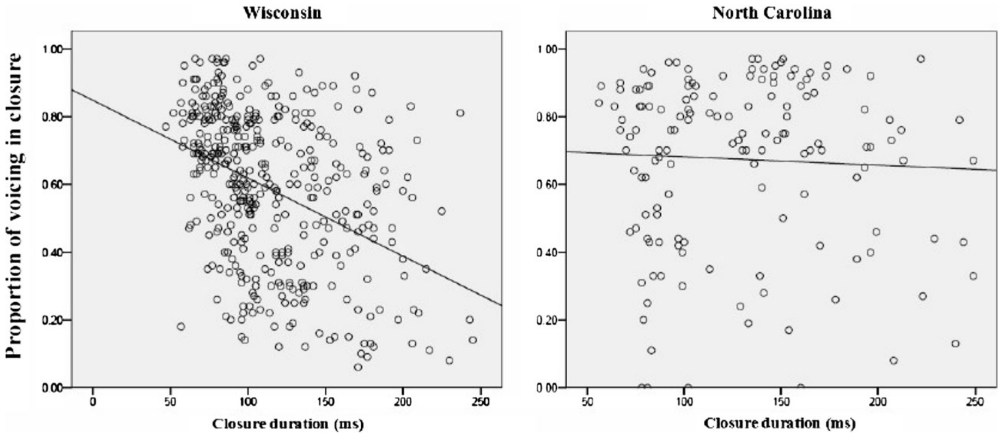

Figure 7 A regression between the proportion of voicing in closure and closure duration for Wisconsin and North Carolina speakers.

the production of the stop was introduced by varying the degree of emphasis of the target word beginning with the stop, vowel quality and the status of voicing of the word-final consonant cluster. Considering all these factors, the results provided an explanation for the impressionistic perception stated at the outset that North Carolina speakers seem to produce more sonorous variants of the stops as compared to Wisconsin speakers.

Comparing the temporal relations between the stop closure and the word, we found on average a slightly longer closure and slightly shorter word for WI speakers. This resulted in a greater proportion of closure-to-word duration for WI than for NC dialect. Also, the WI speakers showed a pattern which was absent in NC productions, namely that the $Prop_{\mathrm{CWD}}$ was greater for words terminating in voiceless consonants than in voiced. Thus, the stop closure was proportionally longer in shorter *b_t* words for WI speakers but not for NC speakers. For each dialect, however, the effects of emphasis were overwhelming. Both the stop closure and word duration were longest in the highest emphasis position, giving rise to the largest $Prop_{\mathrm{CWD}}$. With each position of lesser emphasis, both the stop closure and word durations became progressively shorter, which was also reflected in their progressively smaller $Prop_{\mathrm{CWD}}$. Finally, the effects of vowel category were less consistent and no clear pattern could be detected at present.

Significant dialectal differences were found in the closure voicing. The WI closures were usually not fully voiced and, on average, the voicing portion of the closure did not last longer than 67% of closure duration, which terminated in a complete silence followed by a closure release. Produced by NC speakers, the proportion of closure voicing was much greater, reaching an average of 92%. Further dialectal differences were found in the proportion of closure voicing as a function of word emphasis. For WI speakers, words in high emphasis positions had the smallest proportion of closure voicing whereas the non-emphatic positions yielded the greatest proportion. For NC speakers, the same variation in closure length did not have an effect on the proportion of closure voicing. These differences are illustrated in figure 7, which shows a regression between the proportion of voicing in closure and closure duration.

As can be seen, the general trend for WI speakers indicates that as closure duration increases (as a function of either word emphasis or the word-final voiceless consonant cluster), the proportion of voicing in closure decreases. Linear regression analysis indicated that closure duration was a significant predictor of $Prop_{\mathrm{VC}}$ ($r^2 = .176, p < .001$). This finding is consistent with the predictions stemming from the model of the breath-stream control mechanism (Westbury & Keating 1986) which suggests that a stop occurring in the intervocalic position

Table 1 Summary of inter- and intra-speaker variation in the number of incomplete closure voicing as a function of word emphasis level (High, Intermediate, Low) for North Carolina speakers.

Speaker	High	Intermediate	Low
NC_01	3	3	0
NC_02	15	11	2
NC_03	2	6	1
NC_04	17	9	11
NC_05	6	9	8
NC_06	3	2	0
NC_07	4	10	3
NC_08	0	1	0
NC_09	8	14	1
NC_10	0	0	1
Total	58	65	27

Table 2 Summary of inter- and intra-speaker variation in the number of incomplete closure voicing as a function of word emphasis level (High, Intermediate, Low) for Wisconsin speakers.

Speaker	High	Intermediate	Low
WI_01	20	17	18
WI_02	7	9	0
WI_03	20	20	20
WI_04	19	7	9
WI_05	20	14	18
WI_06	20	20	18
WI_07	18	17	9
WI_08	20	17	20
WI_09	17	8	11
WI_10	17	12	10
Total	178	141	133

should naturally be largely voiced if its closure is short. However, the data from NC speakers cannot be easily explained by this model. Linear regression analysis showed that closure duration was not a significant predictor of $Prop_{VC}$ for NC speakers as indicated by an almost straight regression line fitted to the data ($r^2 = .002, p = .582$). That is, the NC closures could be fully voiced even if the closure duration was long as observed in the high emphasis position of the word.

Perhaps the most dramatic dialectal difference is the number of closures that were fully voiced. The majority of the fully voiced closures in the present sample were produced by NC speakers. For WI speakers, the fully voiced closures were rather sparse and occurred mostly in the positions of lesser emphasis. This is not to say that all speakers of the respective dialects did and will produce these differences at all times. We may reasonably expect considerable inter-speaker variation as well as numerous intra-speaker departures from the observed general patterns in the context of usual variability in speech. To illustrate this point, tables 1 and 2 summarize the individual speaker variation in the number of closures that were not fully voiced. Based on the present findings, we would expect incomplete closure voicing to occur more often in WI productions than in NC. Given that each speaker produced 60 utterances, an ideal WI subject would produce 20 instances of incomplete closure in each word emphasis position (High, Intermediate, Low) and there would be no incomplete closure voicing in the exemplars of an ideal NC speaker. As the tables indicate, this is hardly the case and each speaker exhibits some variation. Ideal (or close to ideal) subjects were very sporadic (compare NC_08 and NC_10 in table 1 and WI_03 and WI_06 in table 2) but the 'outliers' were also rarely found (compare NC_02 and NC_04 in table 1 and WI_02 in table 2). Despite the present inter- and intra-speaker variation, systematic and significant dialectal patterns were evident in statistical assessment of several variables.

Although the present subjects were comparable in terms of their age, gender, and basic demographics, inter- and intra-speaker variation still exists, most likely due to individual speaker characteristics such as laryngeal control and muscle tension, physiologic age and its impact on how well the larynx functions in phonation, motor function and force of articulators involved in the production of linguistic stress, to name a few. It is likely that individual biomechanics rather than social factors play an important role in the type of variation such as the degree of closure voicing. Pursuing this thought, we examined the background questionnaires of the three 'outliers' listed above in the hope to find some common trends since all of them were born and lived all their lives in the geographic areas of our study. Speaker NC_02, a 58-year-old woman, had four years of college education and was a manager who made frequent trips out of North Carolina, about twelve in a year. Speaker NC_04 was 53 years old, had a high-school education, worked as a teacher assistant and made no more than one trip in a year. Speaker WI_02 was 60 years old, had high-school education, was retired and did not report any travel. Based on these reports, it seems that the education level, occupation and mobility do not have a strong effect on this specific dialectal feature of closure voicing. Reinforcing this point, these three women were comparable with the remaining subjects in the present sample in terms of their socio-economic status.

Clearly, these two different patterns of stop closure voicing for NC and WI speakers suggest a possibility of dialectal differences in the way voicing is maintained during the closure by the speech production mechanism. Before we reach this conclusion, however, it is informative to get some insight into possibilities and mechanisms available to the speaker in order to sustain voicing. The following paragraphs will outline such possibilities and will point to the use of active versus passive maneuvers which seem to underlie the dialectal differences found in the production of /b/ in the intersonorant position.

Voicing during the closure of a bilabial stop consonant is produced by vocal folds vibration which is possible only if the transglottal pressure is sufficient to maintain a positive airflow through the glottis from the lungs. At the beginning of the closure, there is a buildup of intraoral pressure behind the constriction, which causes an expansion of the walls of the vocal tract so that the transglottal pressure decreases. In order for the vocal folds to continue vibrating, the volume of the vocal tract must expand so that the transglottal pressure is sufficient to sustain the vocal fold vibration. As Stevens (1998) points out, the buildup of the intraoral pressure is slower for the voiced stop than for the voiceless stop and glottal vibration can continue throughout the closure interval if the volume of the vocal tract undergoes a steady increase. However, the transglottal pressure will be reduced, the manifestation of which is a decreased frequency and weaker amplitude of the glottal pulses during the closure. The gradually decreasing transglottal pressure will cause a gradual decrease of the amplitude of the glottal pulses until a complete cessation of voicing.

There are both active and passive supplementary adjustments appropriate for the expansion of the volume of the vocal tract to sustain voicing. Hayward (2000) lists three active mechanisms which can be used by the speaker. The primary is lowering the larynx and advancing of the tongue root which will expand the pharynx (also see Stevens 1998 for a detailed discussion). The second is an incomplete closure of the velo-pharyngeal port. This partial velic opening, however, cannot be as great as in the production of a nasal consonant. Rather, its function is to allow an appropriate amount of air to leak out in order to prevent the pressure build-up. Finally, slackening the vocal folds to reduce the size of the pressure drop across the glottis is considered an active adjustment although this position can be problematic because the vocal folds may also vibrate more slowly due to the pressure drop without any control on the part of the speaker. Another active adjustment may include raising of the velum as reported in Bell-Berti (1975) and Künzel (1979), among others.

In terms of passive mechanisms, the expansion of the pharyngeal cavity is considered the most common adjustment. In Hayward's (2000: 240–241) words, 'if the walls of the oro-pharyngeal cavity are slack and compliant, then an increase in the air pressure inside will exert a slight outward push, causing the cavity to expand and the pressure inside to lower. The extent to which this is possible depends on place of articulation. Passive expansion is greatest for bilabials, which have the largest oro-pharyngeal cavity and, accordingly, the largest surface area of compliant walls'. Although the reasoning behind positing the pharyngeal cavity enlargement as a passive mechanism is intuitively appealing, Bell-Berti (1975) concludes that the pharyngeal cavity expansion is neither exclusively passive nor exclusively active. In her study, there was much inter-speaker variation and each speaker seemed to use both modes of enlargement, i.e. passive or active, favoring one over the other. This variation most likely represents behavioral differences reflecting some form of speaker idiosyncrasy or else anatomical differences among speakers.

Modeling efforts which attempted to calculate the precise duration of closure voicing suggest that, for a bilabial stop /b/ in the intervocalic position, voicing will be maintained for about 60 ms without active adjustments on the part of the speaker (Westbury & Keating 1986). In order to sustain closure voicing, the following active maneuvers besides the lowering of the larynx have been listed by Westbury & Keating (1986: 151–152): 'contracting the expiratory muscles, decreasing average area of the glottis and /or tension of the vocal folds, decreasing the level of activity in muscles which underlie the walls of the supraglottal cavity, actively enlarging the volume of that cavity, or creating a narrow opening between the posterior pharyngeal wall and soft palate. These maneuvers, occurring singly or in combination, will have their greatest effect on duration of closure voicing when they occur during the closure interval itself, in concert with the rise of pressure above the glottis which naturally accompanies vocal tract occlusion'. A reanalysis of the closure voicing duration in the present cross-dialectal data against the 60 ms reference interval indicates that, on average, WI speakers may maintain 'natural' voicing (in Westbury & Keating's terms) when producing the bilabial stop in the low emphasis position (means were 54 ms in *b_t* words and 62 ms in *b_d* words) and perhaps in the intermediate position (65 ms and 68 ms, respectively) but not in the high emphasis position (79 ms and 87 ms, respectively). NC speakers, on the other hand, seem to have used some active mechanism or a combination of active mechanisms to sustain closure voicing across all three emphasis positions. Their average values for low emphasis were 76 ms in *b_t* words and 74 ms in *b_d* words, for the intermediate – 81 ms and 82 ms, respectively, and for high emphasis position – 110 ms and 112 ms, respectively.

Given that NC closures were mostly fully voiced and the proportion of voicing during the closure was generally insensitive to the variation in closure length as a function of word emphasis, it appears that NC speakers maintained transglottal pressure during the stop closure by some active articulatory maneuvers, which may or may not be the same as for the WI speakers. Figure 8 displays a spectrogram and the corresponding waveform for NC speaker NC_01 (see table 1) whose closures were mostly fully voiced. The display shows an excerpt *the dull baits* from the sentence *Dad said the dull BAITS are best* in which the stop /b/ in *baits* occurs in high emphasis position. Beside the characteristic low-frequency voice bar, there are visible traces of F1 and F2 throughout the closure of the stop /b/. Although not studied in this paper, the closure of the stop /d/ in *dull* displays similar characteristics, i.e. the low-frequency voice bar and traces of the three higher formants running from left to right across the consonantal closure. The waveform above the spectrogram affirms that there was no silent gap during either consonantal closure at any point in time. The presence of the formant traces may indicate an incomplete velo-pharyngeal closure for the NC speaker to facilitate

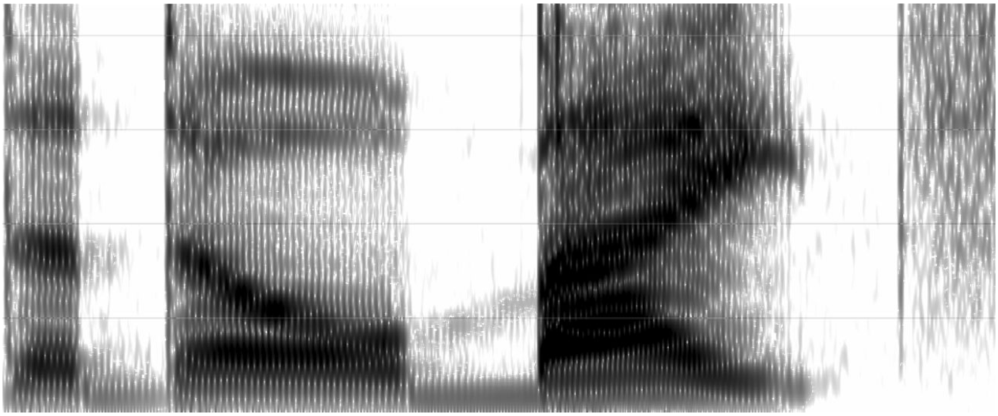

Figure 8 Spectrogram and waveform of the utterance 'the dull baits' in the production of a female speaker from North Carolina. The display is centered on the closure of the stop /b/ when the word 'baits' was highly emphasized.

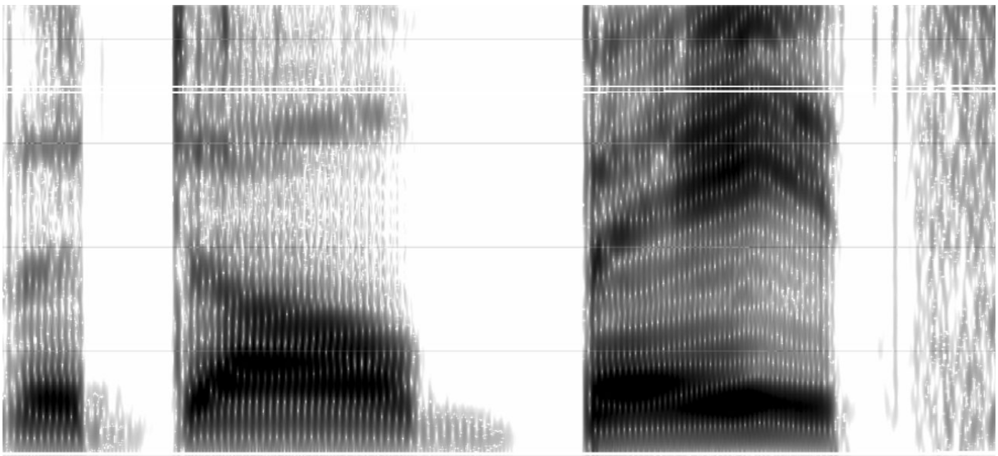

Figure 9 Spectrogram and waveform of the utterance 'the dull baits' in the production of a female speaker from Wisconsin. The display is centered on the closure of the stop /b/ when the word 'baits' was highly emphasized.

the maintenance of voicing. Alternatively, it may suggest some other active maneuvers to sustain voicing although their nature is difficult to determine in the absence of articulatory and aerodynamic measures.

A different way of maintaining voicing during the stop closure is shown in figure 9 for a Wisconsin speaker WI_04 (see table 2). The spectrogram and the waveform display exactly the same utterance as produced by NC speaker. However, as evident from both displays in figure 9, the closures of the WI speaker were not fully voiced. No presence of higher formants can be detected in the closures of either /b/ or /d/ and the low-frequency bar indicating voicing appears to be weaker when compared to that of the NC speaker. Apparently, the expansion of the vocal tract volume was not great enough to maintain the transglottal pressure throughout the entire closure duration. Based on the visual inspection of the spectrograms, it seems unlikely that the WI speaker sustains voicing as actively as the NC speaker does. Again, aerodynamic data are needed in order to confirm this observation. For the sake of a comparison, the closure voicing duration (and the closure duration itself) for the stop /b/ of NC speaker is 132 ms and the voicing portion of the closure of WI speaker is 78 ms which is followed by a 60 ms silent gap. This difference in both timing and the nature of closure voicing suggests involvement of some active mechanism(s) in the production of a voiced stop by NC speaker. However, no active maneuvers can be detected in the production of WI speaker despite the fact that her closure voicing is slightly longer than the 60 ms reference interval for 'natural' voicing proposed in Westbury & Keating's (1986) calculations.

Because this study involves acoustic analysis only and no aerodynamic data are available for the present set of acoustic measurements we cannot assert with certainty that NC speakers utilize the velum to sustain voicing during the stop closure. A possibility of lowering the velum as one probable active mechanism comes from rather informal observations. Namely, the sound quality of the stop itself and of the speech from the majority of our NC speakers in general gives us an indication that the velo-pharyngeal port is at least open partially allowing air to escape through the nasal tract.

It is the case that Appalachian speech has long been described (and stereotyped) as having at least some degree of nasality present even in words that have no nasal segments. This is often called a 'nasal twang' and is commonly found not only in speech of individuals from this area, but in both performed and recorded Appalachian vocal music (especially 'Bluegrass' music). Although we did not complete an acoustic analysis of the presence of nasalization in these Appalachian speakers, from an impressionistic phonetic point of view, we find at least some degree of nasalization present in the speech of almost every NC speaker. Our perception concurs with what Nihalani (1975) reported for voiced stops in Sindhi, instrumentally supporting his proprioceptive impression that voiced stops in Sindhi were slightly nasalized. A very slight opening in the nasal cavity was detected, which 'does not inhibit the formation of a plosive, but it gives to the voiced stop a slight nasalization' (pp. 95–96). A small leakage of air through the nasal cavity came about by an incomplete velo-pharyngeal closure sufficient enough to absorb the transglottal airflow in order to prevent the rise of the transglottal pressure.

Alternatively, lowering the larynx and advancing the tongue root to expand the pharynx will be another possibility for NC speakers to sustain voicing during the closure. In this case, their speech would sound more sonorous, which we have also noted as a feature of this southern dialect. Although it might be argued that an incomplete oral occlusion could be used as another possible active mechanism to sustain voicing, most stops of NC speakers displayed a release burst in the spectrogram which would likely not occur if a complete oral closure were not made.

The present study was limited to investigation of the stop /b/ in the intersonorant position. Although the current experimental design executed some control with regard to the variation of the stop as a function of word emphasis, voicing of the word-final consonant cluster and vowel category in the word, the stimulus material consisted of read sentences. It will be instructive to determine whether these dialectal patterns are maintained in unconstrained informal speech. Since only the bilabial stop /b/ in a specific environment was examined, we do not know whether the present findings can be generalized to all voiced stops in these two dialects. Moreover, we have not addressed the variation in stop production as a function of positional differences in a word and did not examine acoustic characteristics of the stop in the initial and final positions. Finally, we did not inquire into how the contrast between voiced and voiceless stops is maintained in these two regional varieties of American English. These and other questions will be pursued in future research.

Acknowledgements

This work was supported by the research grant No. R01 DC006871 from the National Institute of Deafness and Other Communication Disorders, National Institutes of Health and by an Undergraduate

Research Scholarship from the College of Arts and Sciences at The Ohio State University. Portion of this research was presented at the Acoustical Society of America meeting in Miami, FL, in November 2008. We thank Joseph Salmons and Osamu Fujimura for discussions on the topic. Special thanks go to Adrian Simpson, John Esling and three anonymous reviewers for their comments on an earlier version of this manuscript.

Appendix. Stimulus material used in the present study

/b_ts/ words

bits

John knows the SOFT bits are sharp.
No! John knows the SMALL bits are sharp.

John knows the small SCREWS are sharp.
No! John knows the small BITS are sharp.

John knows the small bits are DULL.
No! John knows the small bits are SHARP.

baits

Dad said the BRIGHT baits are best.
No! Dad said the DULL baits are best.

Dad said the dull HOOKS are best.
No! Dad said the dull BAITS are best.

Dad said the dull baits are WORST.
No! Dad said the dull baits are BEST.

bets

John said the BIG bets are low.
No! John said the SMALL bets are low.

John said the small POTS are low.
No! John said the small BETS are low.

John said the small bets are HIGH.
No! John said the small bets are LOW.

bats

Doc said the LARGE bats are fast.
No! Doc said the SMALL bats are fast.

Doc said the small BIRDS are fast.
No! Doc said the small BATS are fast.

Doc said the small bats are SLOW.
No! Doc said the small bats are FAST.

bites

Sue thinks the LARGE bites are deep.
No! Sue thinks the SMALL bites are deep.

Sue thinks the small CUTS are deep.
No! Sue thinks the small BITES are deep.

Sue thinks the small bites are WIDE.
No! Sue thinks the small bites are DEEP.

/b_dz/ words

bids

Ted thinks the SPRING bids are low.
No! Ted thinks the FALL bids are low.

Ted thinks the fall SALES are low.
No! Ted thinks the fall BIDS are low.

Ted thinks the fall bids are HIGH.
No! Ted thinks the fall bids are LOW.

bades

(The nonsense word *bade* was explained to the speaker as indicating ‘a brand of knife, a brand name’.)

Ted says the SHARP bades are cheap.
No! Ted says the DULL bades are cheap.

Ted says the dull FORKS are cheap.
No! Ted says the dull BADES are cheap.

Ted says the dull bades are WEAK.
No! Ted says the dull bades are CHEAP.

beds

Rob said the SHORT beds are warm.
No! Rob said the TALL beds are warm.

Rob said the tall CHAIRS are warm.
No! Rob said the tall BEDS are warm.

Rob said the tall beds are COLD.
No! Rob said the tall beds are WARM.

bads

(The speaker was told that *bad* refers to ‘an error or mistake’. For example, if someone makes an error, he or she might say ‘my bad’ instead of ‘my mistake’.)

Mike thinks the BIG bads are worse.
No! Mike thinks the SMALL bads are worse.

Mike thinks the small GOODS are worse.
No! Mike thinks the small BADS are worse.

Mike thinks the small bads are BEST.
No! Mike thinks the small bads are WORSE.

bides

(The nonsense word *bide* was explained to the speaker as indicating ‘a small animal, a type of dog’.)

Jane thinks the SHORT bides are cute.
No! Jane thinks the TALL bides are cute.

Jane thinks the small CATS are cute.
No! Jane thinks the small BIDES are cute.

Jane thinks the small bides are GROSS.
No! Jane thinks the small bides are CUTE.

References

Bell-Berti, Fredericka.1975. Control of pharyngeal cavity size for English voiced and voiceless stops. *Journal of the Acoustical Society of America* 57, 456–461.

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, 129–159.

Clopper, Cynthia G., David B. Pisoni & Kenneth de Jong. 2005. Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America* 118, 1661–1676.

Crowther, Court S. & Virginia Mann. 1992. Native language factors affecting use of vocalic cues to final consonant voicing in English. *Journal of the Acoustical Society of America* 92, 711–722.

Docherty, Gerard J. 1992. *The timing of voicing in British English obstruents*. Berlin: Foris.

Flege, James E. & W. S. Brown, Jr. 1982. The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance. *Journal of Phonetics* 10, 335–345.

Hayward, Katrina. 2000. *Experimental phonetics*. Harlow: Pearson.

Jacewicz, Ewa, Robert A. Fox & Joseph C. Salmons. 2006. Prosodic prominence effects on vowels in chain shifts. *Language Variation and Change* 18, 285–316.

Kent, Raymond D. & Kenneth L. Moll. 1969. Vocal-tract characteristics of the stop cognates. *Journal of the Acoustical Society of America* 46, 1549–1555.

Kewley-Port, Diane & Malcolm S. Preston. 1974. Early apical stop production: A voice onset time analysis. *Journal of Phonetics* 2, 195–210.

Kingston, John & Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70, 419–454.

Koenig, Laura L. & Jorge C. Lucero. 2008. Stop consonant voicing and intraoral pressure contours in women and children. *Journal of the Acoustical Society of America* 123, 1077–1088.

Künzel, Hermann J. 1979. Some observations on velar movement in plosives. *Phonetica* 36, 384–404.

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Maldon, MA & Oxford: Blackwell.

Labov, William, Sharon Ash & Charles Boberg. 2006. *Atlas of North American English: Phonetics, phonology and sound change*. Berlin: Mouton de Gruyter.

Labov, William, Malcah Yaeger & Richard Steiner. 1972. *A quantitative study of sound change in progress*. Philadelphia, PA: U.S. Regional Survey.

Lisker, Leigh. 1957. Closure duration and the intervocalic voiced–voiceless distinction in English. *Language* 33, 42–49.

Lisker, Leigh. 1986. 'Voicing' in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 19, 3–11.

Lisker, Leigh & Arthur S. Abramson. 1964. A cross-linguistic study of voicing in initial stops: Acoustical measurements. *Word* 20, 384–422.

Milenkovic, Paul. 2003. TF32 software program. University of Wisconsin–Madison.

Nihalani, Paroo. 1975. Velopharyngeal opening in the formation of voiced stops in Sindhi. *Phonetica* 32, 89–102.

Ohala, John J. & Carol J. Riordan. 1979. Passive vocal tract enlargement during voiced stops. In Jared J. Wolf & Dennis H. Klatt (eds.), *Speech communication papers*, 89–92. New York: Acoustical Society of America.

Ohde, Ralph N. 1984. Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America* 75, 224–230.

Purnell, Thomas C., Joseph C. Salmons & Dilara Tepeli. 2005. German substrate effects in Wisconsin English: Evidence for final fortition. *American Speech* 80, 135–164.

Purnell, Thomas C., Joseph C. Salmons, Dilara Tepeli & Jennifer Mercer. 2005. Structured heterogeneity and change in laryngeal phonetics: Upper Midwestern final obstruents. *Journal of English Linguistics* 33, 307–338.

Smith, Bruce L. 1979. A phonetic analysis of consonantal devoicing in children's speech. *Journal of Child Language* 6, 19–28.

Snoeren, Natalie D., Pierre A. Hallé & Juan Segui. 2006. A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics* 34, 241–268.

Stevens, Kenneth N. 1998. *Acoustic phonetics*. Cambridge, MA: MIT Press.

Studenbaker, Gerald A. 1985. A 'rationalized' arcsine transform. *Journal of Speech and Hearing Research* 28, 455–462.

Svirsky, Mario A., Kenneth N. Stevens, Melanie L. Matthies, Joyce Manzella, Joseph S. Perkell & Reiner Wilhelms-Tricarico. 1997. Tongue surface displacement during bilabial stops. *Journal of the Acoustical Society of America* 102, 562–571.

Thomas, Erik R. 2000. Spectral differences in /ai/ offsets conditioned by the voicing of the following consonant. *Journal of Phonetics* 28, 1–25.

Thomas, Erik R. 2001. *An acoustic analysis of vowel variation in New World English*. Durham, NC: Duke University Press.

Thomas, Erik R. 2002. Instrumental phonetics. In J. K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*,168–200. Maldon, MA & Oxford: Blackwell.

Van Hofwegen, Janneke & Erik R. Thomas. 2009. Velarization of word-initial /l/ in African American English. *Journal of the Acoustical Society of America* 125, 2567.

Westbury, John R. 1983. Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America* 73, 1322–1336.

Westbury, John R. & Patricia A. Keating. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22, 145–166.