Journal of Phonetics 39 (2011) 660–667

Contents lists available at ScienceDirect

Journal of Phonetics

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

An articulatory–aerodynamic approach to stop excrescence

Ryan K. Shosted *

Department of Linguistics, University of Illinois at Urbana-Champaign, 4080 Foreign Language Building, 707 S. Mathews Ave., MC-168, Urbana, IL 61801, United States

ARTICLE INFO

Article history:
Received 28 June 2010
Received in revised form
28 June 2011
Accepted 4 July 2011
Available online 5 August 2011

ABSTRACT

The distinction between underlying and excrescent stops in pairs like 'mints' and 'mince' was convincingly demonstrated by Fourakis and Port (1986). Several subsequent studies have been unable to replicate the result for speakers of American English, or have done so only partially. These studies have largely dealt with the acoustic signal. This study presents an approach to stop excrescence that refers to both the aerodynamics and articulation of the phenomenon. The results confirm and expand on the original findings. Using nasal flow as an indirect measure of velopharyngeal aperture and electropalatography (EPG) to estimate the moment of oral release, the presence of occlusion, as well as the duration of nasal and oral occlusion were measured. Overall contact across the palate was also measured. Disyllabic and monosyllabic tokens with /ns/ and /nts/ in final position were pronounced by four male speakers of American English. Disyllabic tokens could be either stressed or unstressed on the final syllable. In Experiment I, speakers produced tokens in a standard carrier phrase; in Experiment II, they produced one of the items in contrastive focus to its 'homophonous' counterpart, e.g., 'I said *mince* not mints'. Underlying stops were significantly longer than excrescent stops, including in the contrastive-focus condition. A trading relation between nasal and oral stop duration was demonstrated when the stop was excrescent, but not when it was underlying. This suggests that the nasal–oral occlusion in epenethetic stops is divided proportionally between the underlying nasal and excrescent oral stop, but that the durations of the nasal and underlying oral stops are independent.

1. Introduction

Surface realizations of a homorganic nasal-fricative cluster like /ns/ are derived from a simple gestural timing model: if the tongue maintains the oral closure for /n/ after the velopharyngeal port (VP) has closed in anticipation of /s/, a homorganic oral stop results, i.e., /ns/ → [nts]. If the VP closes at the moment of the fricative's oral release (or later), the pressure needed for the stop is vented through the nose, resulting in [ns] (Ohala, 1974). Most research on the production of stop excrescence[1] in nasal-fricative clusters has used the oro-nasal acoustic signal to identify the presence and duration of segmental and sub-segmental events. These typically include the length of the preceding nasal [n], the presence and/or duration of a silent phase (interpreted as the voiceless oral closure of [t]), and the presence and/or duration of a burst transient preceding the fricative [s]. Because of ambiguities inherent in the acoustic signal, Yoo and Blankenship (2003, p. 156) recommend more studies of the phenomenon using articulatory instruments.

The literature on stop excrescence in consonant clusters has been thoroughly reviewed by Warner and Weber (2001). Phonetic approaches to the topic have focused on determining: (a) the similarity of underlying (phonological) and excrescent (surface/phonetic) forms and (b) whether speaker–listeners can distinguish them.

The acoustics of excrescent and underlying stops have been shown to vary significantly between speakers, languages, and dialects, to vary only in restricted phonological contexts, or not to vary at all. Results also vary depending on what was measured or counted (closure length, presence of burst, etc.). Fourakis and Port (1986) showed that while four speakers of American English consistently produced an excrescent [t] in /ns/ sequences, five speakers of South African English did not. For American English speakers, nasals before underlying stops were significantly shorter than nasals before excrescent stops, suggesting that not only are underlying and inserted segments 'articulatorily distinct' (p. 210), but they may influence the production of surrounding segments, as well. Several subsequent studies were unable to replicate these results, or did so only partially (Arvaniti, 2006; Blankenship, 1992; Colavin, 2007; Lee, 1990; Yoo & Blankenship, 2003). Using the TIMIT database, Blankenship (1992) found evidence of excrescent [t] (closure and/or burst) in only 26% of /ns/ tokens; excrescent [t] was no shorter than underlying [t].

* Tel.: +1 217 244 7788.
E-mail address: rshosted@illinois.edu

[1] The phenomenon is also referred to as 'epenthesis', 'intrusion', and 'emergence' (Campbell, 2004, pp. 36–37).

0095-4470/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2011.07.001

Yoo and Blankenship (2003) found underlying [t] to be longer, though the significant result was based on measures of only two of 24 words recorded. Dembowski (2003, p. 1917) reported that more than 95% of both /ns/- and /nts/-words manifested closure and/or burst in American English, with no difference in duration. Warner and Weber (2001, p. 63) showed that word-final /ns/ sequences in Dutch categorically manifested a burst. Colavin (2007), using a method explored by Port and Crawford (1989, p. 263), asked speakers to produce forced-contrast differences between monosyllabic words ending in /ns/ and /nts/. Underlying [t], as a proportion of word-length, was longer than excrescent [t] under contrast (both underlying and excrescent [t] were longer than their non-contrastive counterparts).

Exactly whether and how listeners distinguish /ns/ and /nts/ tokens remains unclear. Barnitz (1974) reported that excrescent [t] was unlikely to be perceived in /ns/ word-medially. Ali, Daniloff, and Hammarberg (1979) found that the majority of /ns/ clusters which listeners identified as having a stop indeed contained a silent gap of ≥ 15 ms and a burst (Ali et al., 1979, p. 95). However, listeners perceived some /ns/ clusters as having an excrescent stop, even though they manifested neither a burst nor a silent gap. In a phoneme-monitoring task (Warner & Weber, 2001), listeners were asked to respond as soon as they heard a [t] during word-final /ns/. Differences in reaction time support the claim that 'epenthetic stops are…perceived differently' (p. 82).

However, when the duration of silent closure and presence of burst were systematically manipulated, Kilpatrick, Shosted, and Arvaniti (2007, p. 656) reported that '[l]isteners have only limited sensitivity to the differences between epenthetic and underlying [t]…' Most responses were not significantly different from chance, even when closure duration was artificially lengthened and a burst transient was inserted or deleted (Kilpatrick et al., 2007, p. 655; Table 2). Some tokens with no evidence of [t] were significantly identified as /nts/ tokens.

Ohala (1974) observes that nasal flow (N_F) in /ms/ subsides when voicing ceases. The absence of N_F (indicative of VP closure) may also suggest that oro-pharyngeal pressure (O_P) is rising (assuming constant subglottal pressure). Mistiming between VP closure and oral opening may reasonably result in a longer silent gap and a stronger burst.

Ali et al. (1979, p. 85) confirmed that '[i]ntrusive stops appear to result from a prolonged oral occlusion of the nasal stop which is released with a vigorous burst release just prior to complete formation of the following fricative slit constriction.' In addition, their results show evidence of nasal frication[2] occurring in up to 1/3 the duration of the fricative noise. Ali et al. (1979) conclude that the 'delayed release of the oral occlusion' delays the production of the oral fricative and point out that during the oral closure the velum may remain open (p. 95). They argue that while epenthesis in /ns/ clusters is natural, intrinsic, and allophonic, it can also be subjected to 'voluntary control' (p. 96; cf. Tatham, 1969).

It has been hypothesized that stress may affect excrescence (Arvaniti, 2006; Blankenship, 1992; Clements, 1987; Yoo & Blankenship, 2003): in polysyllabic words, excrescent [t] may be longer after a stressed than a stressless vowel, but results have been mixed. More generally, the 'articulatory expansion' or 'localized hyperarticulation' of stress and focus[3] is well-researched (Beckman, Edwards, & Fletcher, 1992; Bombien, Mooshammer, Hoole, Rathcke, & Kühnert, 2007; Cho, 2006; Cho & Keating, 2009; de Jong, 1995, 2004; de Jong & Zawaydeh, 2002; Erickson, 2002; Harrington, Beckman, Fletcher, & Palethorpe, 1998; Mooshammer & Fuchs, 2002). Gestures have been shown to be larger, longer, and faster in accented syllables (Bombien et al., 2007; Cho, 2006). However, Cho and Keating (2009, p. 473) found that their EPG measure of maximum contact was not affected by stress or accent. Closure duration, on the other hand, was longer for primary vs. secondary stress and focused vs. unfocused items. These results may have implications for the production of excrescent stops produced under stress and/or focus.

While it is true that excrescent [t] in /ns/ clusters has acoustic and perceptual implications, it is presumed to result from the coordination of VP and oral opening gestures. Articulatory evidence of both gestures may contribute to a clearer picture of the phenomenon. I will present direct evidence of oral opening using EPG and indirect evidence of VP opening using N_F. Data annotation and resulting measures rely most heavily on the N_F and EPG signals.

2. Methods

2.1. Instrumentation

Data were recorded using the Articulate Instruments (AI) software interface (v. 1.17; Articulate Instruments, Musselburgh, UK). Each speaker was recorded while wearing a Reading (62-sensor) electropalate (Hardcastle, Jones, Knight, Trudgeon, & Calder, 1989) manufactured by Incidental (Newbury, Berkshire, UK). Each electropalate was designed using a unique superior maxillary model cast by a board-certified orthodontist. Secured in the subject's mouth, the palate was connected to the WinEPG EPG3 Serial Interface SPI V2.0. Whenever the tongue touches an electrode, an electrical signal registers. Given that the position of each electrode is known, the electrode signals are ultimately interpreted as indicating whether a specific region of the palate is in contact with the tongue (Baken, 1987; Hardcastle, 1972). The sampling rate of the palate scanner is 100 Hz.

Speakers wore a head-mounted AKG C520 cardioid microphone (Harman International, Stamford, CT). The audio signal was passed to a Grace m101 pre-amplifier (Grace Design, Boulder, CO) and then to one of two line inputs of the SPI.

Subjects also wore a Gold Seal nasal CPAP mask (Philips Respironics, Eindhoven, The Netherlands). The vent of the nasal mask was plugged with a heated Fleisch Pneumotach oriented so that air from the mask was channeled through the pneumotach (Baken, 1987; Fleisch, 1925) (Biopac TSD137; Biopac Systems, Goleta, CA). The pneumotach pressure ports were connected via rubber cannulae to a Biopac TSD160B pressure transducer (± 12.5 cm H_2O). The voltage output of the transducer was amplified by a differential bridge amplifier (gain=5000). This signal was then passed to the other line input of the SPI. This signal was calibrated using a positive/negative-discharge gas pump.

From the SPI, the audio and aerodynamic signals were passed to an AudioFire2 (Echo Digital Audio, Carpinteria, CA) IEEE 1394 serial bus interface for isochronous real-time data transfer to an HP xw4400 Workstation running Microsoft Windows XP (Version 2002, Service Pack 3). The EPG signal was passed from the SPI to the same computer via USB. Synchronization of the audio, aerodynamic, and EPG signals is a function of the WinEPG hardware and Articulate Instruments Ltd. (2008) 1.17 software. The AudioFire2 interface was used to transmit the audio and aerodynamic signals directly to the computer. Synchronization was verified, and in some cases rectified, by comparing the EPG and audio signals for a word-initial /t/ (recorded among the distractor items for each subject). The author verified that the acoustic excitation (burst/aspiration) of the word-initial /t/ (in the acoustic waveform) coincided with the release of occlusion in the EPG signal. The time difference between these events, if any, was used to

[2] The airstream becomes turbulent at a point in (or on the way out of) the nasal cavity.

[3] In this study I use de Jong and Zawaydeh's (2002, p. 54) most basic formulation of phonological focus as stress that is not restricted to the syllable.

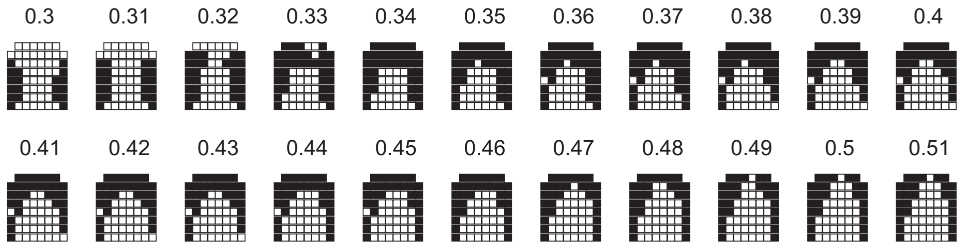

Fig. 1. Palatograms for alveolar cluster in *prints* (cf. Fig. 3). Each palate is numbered according to its time stamp (in seconds). Black squares indicate electrodes in contact with the tongue. Complete occlusion occurs from 0.33 to 0.48 s.

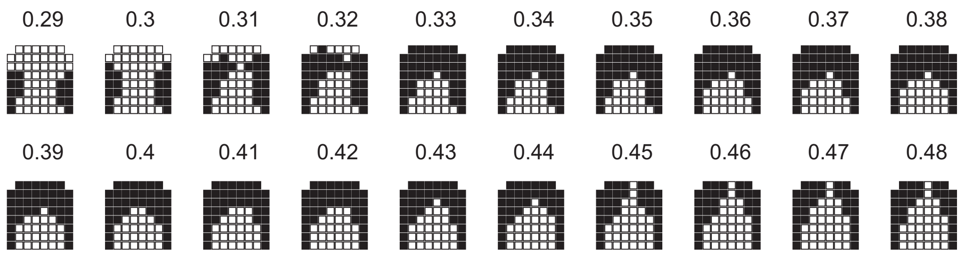

Fig. 2. Palatograms for alveolar cluster in *prince* (cf. Fig. 4). Conventions as in Fig. 1. Complete occlusion occurs from 0.32 to 0.44 s.

synchronize the signals. N_F was captured in line with the audio, so synchrony of N_F and EPG signals was dependent only on the synchrony of the audio and EPG. Audio and N_F were digitally sampled at 11,025 Hz.

2.2. Speakers

Four male speakers of American English between the ages of 21 and 33 participated in the study. The subjects were from Washington (Speaker 1), California (Speaker 2), Utah (Speaker 3), and North Carolina (Speaker 4). The author was one of the speakers (Speaker 3).[4]

2.3. Materials

Tokens with word-final /ns/ and /nts/ were chosen from materials[5] used in previous experiments (Ali et al., 1979; Kilpatrick et al., 2007; Yoo & Blankenship, 2003); there were six monosyllabic items and eight disyllabic items: mince, quince, prince, mints, quints, prints; ˈoffense, ˈdefense, ˈcontents, ˈpresents, ofˈfense, deˈfense, conˈtents, preˈsents. To reduce variance in the measures, monosyllabic pairs were selected because they shared a vowel, /ɪ/; pair members differed only in the presence/absence of underlying /t/. Disyllabic tokens (three /ns/ and three /nts/ items) all shared the second vowel /ɛ/ (under stress); pair members differed only in stress placement.

2.4. Production tasks

Speakers participated in two experiments. The first (Experiment I) was a standard production task and the other (Experiment II) was a forced-contrast task or 'dictation sentence' task (Port & Crawford, 1989, p. 263). In Experiment I, speakers uttered monosyllabic and disyllabic tokens in randomized order, interspersed with eight pairs of homophones like *flour* and *flower* to distract the speakers from the presence or absence of an orthographic ⟨*t*⟩ in the test materials. They read tokens in a carrier phrase, e.g., 'Say mints again'; the list (including distractors) was recorded three times. In Experiment II, also repeated three times, speakers uttered tokens in a carrier phrase along with a contrastive token (Colavin, 2007; Port & Crawford, 1989). For monosyllables, the contrastive token differed in terms of underlying-/t/ status; for disyllables, it differed in terms of stress placement only. The token and its counterpart were produced in a carrier phrase, e.g., 'I said *mince* not mints'. Speakers were instructed to place emphasis on the first item (the 'focused' item). The researcher modeled the desired intonation using a few non-test phrases. Speakers were instructed to produce the contrast sentence as if a real instance of ambiguity had occurred in conversation and the speaker wished to distinguish the two words. Eight distractors of the form 'I said *flour* not flower' were interspersed throughout the task. Each test token appeared in focused and unfocused position, doubling the total number of tokens for the forced-contrast task.

In Experiment I, $6\times3\times4=72$ test monosyllables and $8\times3\times4=96$ test disyllables were produced. In Experiment II, 72 test monosyllables were produced under focus and 72 test monosyllables were produced without focus; likewise, 96 test disyllables were produced under focus and another 96 were not.

2.5. Post-processing

Data were converted to a Matlab structure combining the synchronized audio, N_F, and EPG signals. A zero-phase distortion digital filter (lowpass cutoff = 1 kHz) was applied to the N_F data using the *filtfilt* function. A graphical user interface was designed for simultaneous visualization of all three signals. A script was written to detect the presence of complete horizontal occlusion across the palate (Cho & Keating, 2009). The result was a binary signal (0 for absence of complete occlusion, 1 for presence of complete occlusion). The signal can be seen in the top panel of Figs. 3 and 4 where it is labeled Occlusion$_H$; the associated raw palatograms can be seen in Figs. 1 and 2, respectively.

2.6. Annotation protocol and measurement

The phonetic initiation and termination of [n] and [t] were determined based on N_F and EPG occlusion. The acoustics were

[4] The inclusion of the author as a research subject should be treated with caution as self-awareness may have contributed to hyperarticulation.

[5] Materials were presented in orthographic form, considering research suggesting that orthography may be 'mandatorily activated in speech production by literate speakers' (Damian & Bowers, 2003). Spelling pronunciation, therefore, cannot be ruled out.

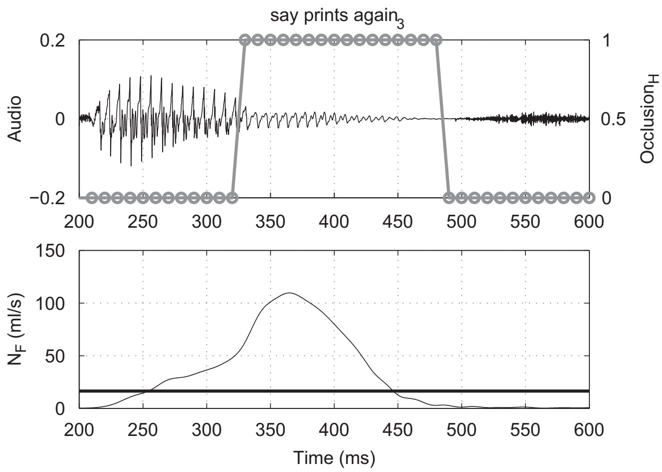

Fig. 3. Audio, EPG occlusion, and N_F signals for one instance of *prints* uttered by Speaker 4. The signal shows /ɪnts/. Corresponding raw palatograms appear in Fig. 1. Complete occlusion occurs from 330 to 480 ms.

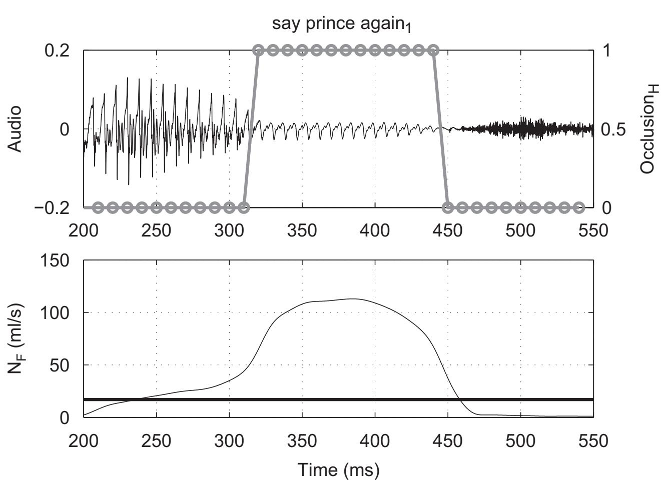

Fig. 4. Audio, EPG occlusion, and N_F signals for one instance of *prince* uttered by Speaker 4. The signal shows /ɪns/. The raw palatograms appear in Fig. 2. Complete occlusion occurs from 320 to 440 ms.

used to locate the general vicinity of /n(t)s/, the boundaries of which (on the first pass) could include part of the preceding vowel and the fricative. A script was written to detect the following points in this vicinity: (1) The first point at which total linguopalatal occlusion occurs while N_F is positive (this point is interpreted as the beginning of the nasal stop). (2) During the linguopalatal occlusion, the first point at which N_F falls to 15% (see below for justification) of its occlusive maximum (this point is interpreted as the beginning of the oral stop). (3) The last point at which total linguopalatal occlusion is recorded (this is interpreted as the end of [t]). If N_F did not fall to 15% of its maximum during the occlusion, this was interpreted as the absence of [t]. The outcomes of some annotations are shown in Figs. 3 and 4. Distances between the annotated points were calculated. General absence of occlusion (including deletion of /n/) was tabulated for relevant tokens. The maximum percentage of linguopalatal contact during the occlusion (including both [n] and [t], if present) was measured automatically (Fontdevila, Pallarès, & Recasens, 1994).

The determination of the 15% threshold in N_F deserves further comment. An absolute threshold seemed inappropriate for determining the onset of the oral consonant since the absolute value of N_F varies with subglottal pressure and therefore potentially with speaker, token, and repetition. A proportional value seemed better suited since it is tied to the N_F of the preceding nasal and thus normalizes for varying subglottal pressure. The 15% threshold was selected using aerodynamic data gathered by Dotevall, Lohmander-Agerskov, Almquist, and Bake (1998) for comparable /nt/ clusters in conjunction with a physiological-aerodynamic model of the oro-nasal tract (Moon & Weinberg, 1985; Warren & Dubois, 1964). It was reasoned that a threshold below 15% would underestimate the occurrence and length of oral consonants by categorizing VP openings $<5\ \text{mm}^2$ as nasal, when VP openings of this size are known to occur even during oral stops produced by normal speakers (Warren, Dalston, & Mayo, 1993). After the initial drop in N_F, successive fluctuations above the threshold (if they occurred) were not considered continuations of the preceding nasal stop. Variations in velar height can yield small changes in flow even once the VP is closed (or nearly closed); moreover, slow variation is detectable even in smoothed digital signals.

2.7. Statistical analysis

Less than 3% (14/504) of the tokens in the corpus manifested no occlusion (oral or nasal); these were removed from further analysis. For these unoccluded tokens, the channel width (the narrowest gap between contact on either side of the palate) during nasalization ranged from one to four sensors, with an average channel width of 2.3 sensors.

R 2.8.1 was used to analyze duration of [t] (t-dur), the proportional duration (Prop-dur) of [t] with respect to the occlusion ($\text{dur}_t/(\text{dur}_n+\text{dur}_t)$), and the maximum percentage of linguopalatal contact during the entire occlusion, if present. Contingency tables were created using R's *xtabs* function. Pearson's χ^2-test for count data (*chisq.test*) was applied to the contingency tables. For each speaker, duration and maximum contact measures over the three repetitions of each token were averaged (tokens with no oral occlusion were excluded from the means). Pearson's product moment correlation coefficient was calculated for the relationship between the logarithm of mean nasal and oral stop durations (data were transformed to reduce right-skew in the distribution of durations). Means were incorporated in linear mixed effects models using the *lme* function in R's *nlme* package. For Experiment I, underlying-/t/ status was included as a fixed effect and speaker was included as a random effect, with token as a nested random effect (Baayen, Davidson, & Bates, 2008; Pinheiro & Bates, 2000). For Experiment II, focus and stressed syllable (1st or 2nd) were included as additional fixed effects.

3. Results

3.1. Oral stop duration

Tokens with an underlying /t/ manifested a longer oral occlusion than tokens without. In Experiment I, [t] was significantly longer for /nts/ tokens both for monosyllables [$F(1,19)=7.76$, $p<0.05$: 20 ms (SD=8) for /ns/; 32 ms (SD=12) for /nts/] and disyllables [$F(1,25)=15.34$, $p<0.001$: 17 ms (SD=9) for /ns/; 34 ms (SD=17) for /nts/]. For disyllables in Experiment I, stressed syllable was associated with [t]-duration, as well, with longer oral stops occurring when stress was on the second syllable [$F(1,25)=7.24$, $p<0.05$: 20 ms (SD=10) for stress on 1st syllable; 32 ms (SD=19) for stress on 2nd]. In Experiment II, for monosyllables, /nts/ tokens manifested longer oral stops than /ns/ tokens [$F(1,19)=20.74$, $p<0.001$: 35 ms (SD=21) for /ns/; 86 ms (SD=50) for /nts/]. For disyllables, underlying-/t/ status was significantly associated with [t]-duration [$F(1,26)=12.63$, $p<0.01$: 35 ms (SD=20) for /ns/; 54 ms (SD=32) for /nts/]; stressed syllable was not significantly associated with [t]-duration.

In Experiment I, proportional duration of [t] was associated with underlying-/t/ status for monosyllabic tokens [$F(1,19)=4.54$, $p<0.05$: 0.22 (SD=0.11) for /ns/; 0.30 (SD=0.12) for /nts/] (Fig. 5). This was also true of disyllabic tokens [$F(1,26)=9.74$, $p<0.01$: 0.19 (SD=0.12) for /ns/; 0.32 (SD=0.13) for /nts/] (Fig. 6).

In Experiment II, proportional duration of [t] was associated with underlying-/t/ status for monosyllabic tokens [$F(1,19)=17.29$, $p<0.001$: 0.31 (SD=0.19) for /ns/; 0.49 (SD=0.17) for /nts/] (Fig. 7). The same was true of disyllabic tokens [$F(1,26)=11.96$, $p<0.01$: 0.30 (SD=0.15) for /ns/; 0.41 (SD=0.19) for /nts/] (Fig. 8). Neither stress nor focus were associated with the proportional duration of [t] in Experiment II.

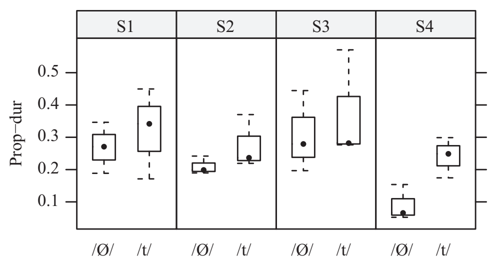

Fig. 5. Proportional duration of [t] in monosyllables with and without underlying /t/, grouped by speaker (Experiment I).

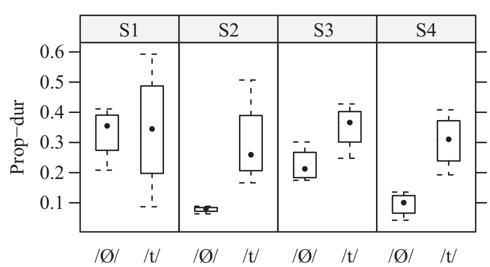

Fig. 6. Proportional duration of [t] in disyllables with and without underlying /t/, grouped by speaker (Experiment I).

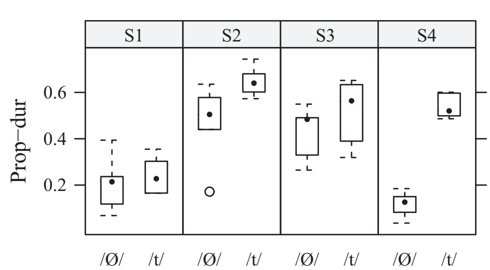

Fig. 7. Proportional duration of [t] in monosyllables with and without underlying /t/, grouped by speaker (Experiment II).

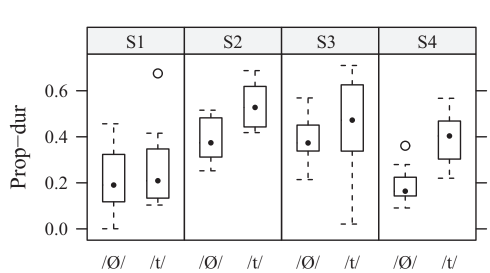

Fig. 8. Proportional duration of [t] in disyllables with and without underlying /t/, grouped by speaker (Experiment II).

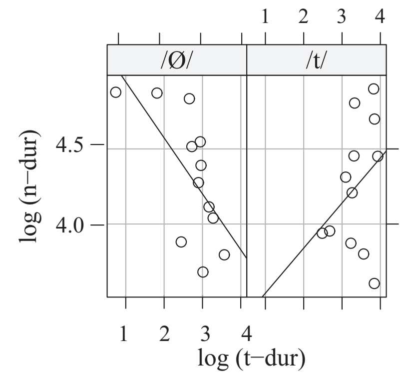

Fig. 9. Monosyllables with and without underlying /t/ in Experiment I: Log of nasal duration (n-dur) plotted against log of oral alveolar stop duration (t-dur) with regression lines. Tokens where [t]-duration=0 have been removed. The correlation was only significant ($p<0.05$) for words without underlying /t/ (left panel).

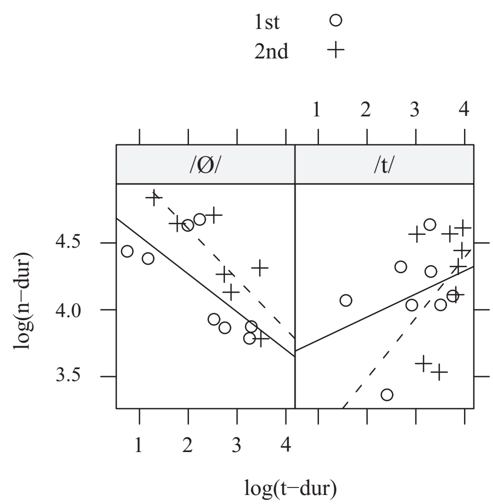

Fig. 10. Disyllables with and without underlying /t/ grouped by stressed syllable in Experiment I: Log of nasal duration (n-dur) plotted against log of oral alveolar stop duration (t-dur) with regression lines (dashed=stress on second syllable). Tokens where [t]-duration=0 have been removed. The correlation was only significant ($p<0.05$) for words without underlying /t/ which also had stress on the second syllable (dashed line, left panel).

3.2. *Correlations between nasal and oral stop duration*

In Experiment I, monosyllables without underlying /t/ (e.g., prince) showed a significant inverse relationship between nasal and oral stop duration [$r=-0.66$, $p<0.05$]; monosyllables with underlying /t/ did not [$r=0.33$, $p>0.05$] (Fig. 9). In Experiment I, disyllables without underlying /t/ with stress on the second syllable (e.g., de'fense) also showed a significant inverse relationship between nasal and oral stop duration [$r=-0.83$, $p<0.05$] (Fig. 10).

In Experiment II, focused monosyllables without underlying /t/ (e.g., in 'I said *prince*…') showed a significant inverse relationship between nasal and oral stop duration [$r=-0.87$, $p<0.001$] (Fig. 11). No significant relationship between these variables was discovered for disyllables in Experiment II.

3.3. *χ^2 results*

Frequency of occlusion was calculated to determine whether or not any of the fixed effects (underlying-/t/ status, stressed syllable, or focus) were associated with the absence of occlusion generally or the absence of [t]-occlusion in particular. Counts of [t]-occlusion

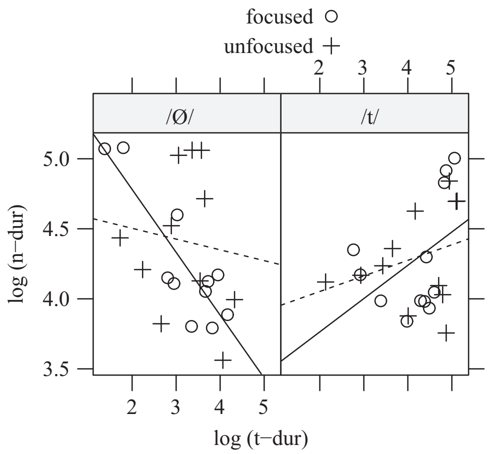

Fig. 11. Monosyllables with and without underlying /t/ grouped by focus in Experiment II: Log nasal duration (n-dur) plotted against log of oral alveolar stop duration (t-dur) with regression lines (dashed=unfocused). Tokens where [t]-duration=0 have been removed. The correlation was only significant ($p<0.001$) for focused monosyllables without underlying /t/ (solid line, left panel).

Table 1
Count data for [t]-occlusions (t-Occ$_1$) and non-occlusions (t-Occ$_0$) produced during Experiment I. The independent variable is presence or absence of underlying /t/. Tokens with no occlusion (including /n/) have been removed.

Non-contrastive						
	σ		ˈ$\sigma\sigma$		σˈσ	
	t-Occ$_0$	t-Occ$_1$	t-Occ$_0$	t-Occ$_1$	t-Occ$_0$	t-Occ$_1$
/∅/	6	28	6	15	6	18
/t/	1	35	3	20	2	22
χ^2	2.80, $p=0.09$		0.81, $p=0.37$		1.35, $p=0.25$	

Table 2
Count data for [t]-occlusions (t-Occ$_1$) and non-occlusions (t-Occ$_0$) produced during Experiment II (focused condition only). The independent variable is presence or absence of underlying /t/. Tokens with no occlusion (including /n/) have been removed.

Contrastive: focused				
	σ		$\sigma\sigma$	
	t-Occ$_0$	t-Occ$_1$	t-Occ$_0$	t-Occ$_1$
/∅/	7	26	12	36
/t/	1	35	3	42
χ^2	4.05, $p=0.04$		4.50, $p=0.03$	

and accompanying χ^2-test results are found in Tables 1 (Experiment I; no significant results), and 2 (Experiment II).

General absence of occlusion (including /n/) was not common (it occurred in less than 3% of the corpus) and it was not associated with /ns/. However, absence of oral occlusion was associated with monosyllabic /ns/ tokens under focus [$\chi^2(69)=4.05$, $p<0.05$] (Table 2, left). This was also true of disyllabic /ns/ tokens under focus [$\chi^2(93)=4.50$, $p<0.05$] (Table 2, right). In other words, when speakers forced the distinction between /ns/ and /nts/ tokens, there was a tendency to realize [t] for /nts/ items and to suppress [t] for /ns/ items. Stressed syllable did not have a comparable effect on the production of nasal-fricative clusters.

3.4. Linguopalatal contact

Percentage of palatal contact during the occlusion was not associated with any independent variable in either Experiment I or II.

4. Discussion

4.1. Comments on the method

The present study combines EPG and aerodynamics in a novel way to examine stop excrescence from an articulatory angle. A critical assessment of this method is warranted. Oral closure has been operationalized here as the presence of linguopalatal occlusion accompanied by nasal flow at or below 15% the maximum value observed during [n]. Adhering to current convention (Cho & Keating, 2009, p. 471), the release of oral occlusion has been defined quite narrowly as the deactivation of at least one electrode that results in an unobstructed channel to the front of the palate.[6] Combined, these conventions suggest that some VP opening is possible during oral stops but oral stops cannot manifest any oral opening without crossing an ontological boundary. Indeed, oral stops may manifest a small degree of velic leakage without being perceived as nasal (Warren et al., 1993), but it is not clear how much oral leakage a stop may manifest without being perceived as a fricative or an approximant. Given current knowledge, the low spatial resolution of the Reading palate and relatively low sampling rate of the palate scanner would make it difficult to accurately determine the dimensions of the opening based on the contact pattern alone. Oral flow, linguopalatal contact, and vertical tongue position could be measured in tandem to model the minimal dimensions of oral opening in a stop release. For now, the first sign of deactivation (whether of one or multiple sensors simultaneously) after occlusion is taken as the best approximation of oral opening.

In addition, the WinEPG system's relatively low sample rate should lead to some skepticism of events whose duration is found to be ≤ 10 ms. To the extent possible, this problem has been remediated by integrating higher-sampled nasal flow into the duration measures. Improvements in EPG technology or use of articulatory technology with a higher sampling rate may yield more accurate results.

Finally, the use of nasal flow as an indirect measure of VP opening is not ideal and should be interpreted with caution. However, direct VP imaging methods like fiberoscopy (Amelot, 2004) and MRI (Story, Titze, & Hoffman, 1996) might also require some form of thresholding to determine effective VP closure since trace amounts of velic leakage are likely even in oral consonants (Warren et al., 1993).

4.2. Oral stop duration

Fourakis and Port (1986) observed that in American English, excrescent stops are shorter than underlying stops. This has been confirmed in the current study, both in Experiment I (standard production) and Experiment II (contrastive-focus). In Experiment I, the average difference between [t] in /ns/ and /nts/ was as great as 17 ms (disyllables), and in Experiment II, the average difference was as great as 51 ms (monosyllables). The duration difference in Experiment I is somewhat larger than the /ns/-/nts/ difference (average ≈ 8.5 ms) reported in Fourakis and Port (1986, p. 207, Table III) but comparable to the average difference (≈ 20 ms) reported in Yoo and Blankenship (2003, p. 161, Figs. 3 & 4). As a proportion of total occlusion duration, the oral stop was

[6] It is possible that de-occlusion may actually manifest itself by a number of simultaneous deactivations.

significantly longer when underlying. Speaker differences were observed. For example, in Fig. 6, the duration of [t] in disyllables is systematically longer in words with underlying /t/ for three out of four speakers, while for one speaker (S1) the opposite may be true.

In the present results, the durations of [n] and [t] are inversely correlated only for items with excrescent [t]. In Experiment I, this is true of monosyllabic /ns/ items and disyllabic /ns/ items (with stress on the second syllable, e.g., de'fense). In Experiment II, this is true of monosyllabic focused items, e.g., ‘I said *mince*…’. According to Fourakis and Port’s (1986) /ns/-/nts/ data, there seems to be an indirect relation between the duration of [n] and [t] regardless of underlying-/t/ status. While the authors do not specifically argue for a trading relation between the length of [n] and [t] in the case of clusters containing the underlying vs. excrescent stop, the present results may suggest such a relation.

This suggests that not only does underlying-/t/ status predict the duration of [t], there is also a basic timing unit for the oral closure which is divided proportionally between nasal and oral stops whenever excrescence occurs in /ns/. When both /n/ and /t/ are present phonemically, their durations are unrelated to each other.

4.3. The presence of occlusion and the presence of [t]

General absence of occlusion (including /n/-deletion) was not common and it was not associated with /ns/ tokens. Though it is perhaps not widespread enough to be a confounding factor in acoustic studies of stop excrescence, general de-occlusion should not be discounted as an articulatory strategy in sequences of nasal and voiceless fricative (see Ohala & Busà, 1995 for a perceptual account of the phenomenon).

Unlike general absence of occlusion, absence of an oral stop was significantly associated with underlying-/t/ status for both monosyllabic and disyllabic items, though only under contrastive focus (Experiment II). The occurrence of the excrescent stop was frequently suppressed during /ns/ tokens. This suggests that VP closing and oral opening can be timed with some precision under the effects of hyperarticulation, supporting the claim that speakers can voluntarily control excrescence (Ali et al., 1979, p. 96). This also highlights the finding that, under non-contrastive conditions, /ns/ tokens are not necessarily distinguished from /nts/ tokens by the complete absence of closure, as was the case (categorically) for Fourakis and Port’s South African English speakers.

4.4. Linguopalatal contact

Degree of linguopalatal contact during the occlusive phase of /n(t)s/ was not associated with underlying-/t/ status, nor with stress or focus. This is comparable to the null effect of stress and focus found for linguopalatal contact in Cho and Keating (2009). Presently, there is no evidence that the occlusions in /ns/ and /nts/ clusters differ systematically in terms of overall contact across the palate.

4.5. Effects of focus and stress

In Experiment II, stop durations in both /ns/ and /nts/ were generally longer than they were in Experiment I. This is comparable to findings that segments are longer in lexically stressed syllables (Klatt, 1976; Lehiste, 1970; Mooshammer & Fuchs, 2002), in (pitch-) accented syllables (Cambier-Langeveld, 1999; Cho & Keating, 2009; de Jong, 1995), and accented words (Turk & White, 1999). However, neither [t]-duration nor proportional duration of [t] were significantly longer for focused vs. unfocused items in Experiment II. This is presumably because /ns/ and /nts/ items were balanced and excrescent oral stops were not lengthened simply because they appeared in focused position. In other words, ‘prince’ does not manifest a longer epenthetic [t] if it is contrasted with ‘prints’. The focus condition did not eliminate the general result of Experiment I, viz., that underlying stops are longer than excrescent stops (Figs. 7 and 8). More frequent suppression of excrescent [t] was observed for /ns/ tokens under focus (Table 2). If anything different happens when /ns/ and /nts/ are contrasted (besides general lengthening of both underlying and epenthetic [t]), it appears to be suppression of [t] in /ns/ sequences.

Yoo and Blankenship (pp. 161–162) commented on the possibility that stress placement could lead to resyllabification in the case of word-medial /ns/ and /nts/ clusters, thus influencing duration. While only word-final clusters are examined in the present study, stress was found to be a significant factor for [t]-duration in Experiment I. On the second syllable (where the /ns/ or /nts/ occurred for all test items), stress significantly increased the duration of the oral stop. Unlike focus, stress has a more general effect, lengthening phonetic [t] regardless of its phonological status.

5. Conclusion

The results of the present study, based on articulatory and aerodynamic measures, are generally supportive of Fourakis and Port (1986), a study based on acoustic measures. It is not clear why these results have not been replicated in subsequent acoustic studies of American English (Blankenship, 1992; Dembowski, 2003; Lee, 1990) or have been replicated only partially in others (Arvaniti, 2006; Colavin, 2007; Yoo & Blankenship, 2003). Warner and Weber (2001, p. 57) note that discrepancies may arise from variability in the data, at least for the corpus-based studies. Yoo and Blankenship (2003, p. 156) note limitations of using oro-nasal sound pressure as the sole basis for annotation in studies of nasal-fricative stop excrescence.

Using articulatory and aerodynamic methods, this study generally supports the findings of Fourakis and Port (1986) regarding stop excrescence in /ns/ clusters. Durational cues of /t/ are found to be more robust for distinguishing /ns/ and /nts/ items than are the presence or absence of closure. Only when consciously contrasting /ns/ and /nts/ pairs do American English speakers behave like the South African English Speakers in Fourakis and Port (1986), reducing the occurrence of closure in /ns/ items (though not categorically, as in South African English). Speakers appear to time VP opening and oral opening more systematically under the effects of contrastive focus. This suggests that speakers can voluntarily control excrescence, as hypothesized by Ali et al. (1979), and may be indicative of fine-grained VP control (Bell-Berti, 1993). In the case of forced-contrast, speakers do not systematically lengthen the oral stop in focused items, but continue to produce a systematic difference between underlying (long) and excrescent (short) oral stops. However, when a (disyllabic) token manifests stress on the second syllable (where /ns/ and /nts/ always occur in these materials), the oral stop may be lengthened, regardless of its underlying status.

The findings of Fourakis and Port (1986) are further extended by suggesting a trading relation between [n] and excrescent [t]. Longer excrescent stops seem to result in shorter nasal stops, whereas the length of underlying /t/ is independent of [n]-duration. By suggesting a different timing strategy for the two sequences, this finding further supports the contention that excrescent and underlying /t/ are articulatorily distinct and that, in American English, the neutralization of /ns/ and /nts/ is still incomplete.

Acknowledgments

I gratefully acknowledge D.H. Whalen and three anonymous reviewers for comments on various drafts of this paper; Dr. Mark Huls and his staff for casting the dental models; and John Broughton and Bertha Bates (InciDental) for manufacturing the EPG palates. I also thank Alan Wrench, Rebecca Colavin, Megan Osfar, Ann Bradlow, Jennifer Cole, José Ignacio Hualde, Chilin Shih, and Christopher Carignan for their insights. The contributions of these individuals in no way changes my responsibility for any errors and omissions that remain. This project was funded in part by Illinois Research Board Grant #08057.

References

Ali, L., Daniloff, R., & Hammarberg, R. (1979). Intrusive stops in nasal-fricative clusters: An aerodynamic and acoustic investigation. *Phonetica*, *36*, 85–97.

Amelot, A. (2004). *Étude aérodynamique, fibroscopique, acoustique et perceptive des voyelles nasales du français*. Ph.D. Thesis. Sorbonne Nouvelle: Université Paris III.

Articulate Instruments Ltd. (2008). *WinEPG Installation and Users Manual*. (Revision 1.16 ed.). Musselburgh, UK: Articulate Instruments Ltd.

Arvaniti, A. (2006). *Stop epenthesis revisited*. Paper presented at LabPhon 10. Abstract available online at 〈http://aune.lpl.univ-aix.fr/~labphon10/abstracts/6.pdf〉. Downloaded June 12, 2010.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Baken, R. J. (1987). *Clinical measurement of speech and voice*. Boston, MA: College-Hill Press.

Barnitz, J. G. (1974). Bloom-p-field, Chom-p-sky, and phonetic epen-t-thesis. *Studies in the Linguistic Sciences*, *4*, 1–13.

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 Vol. II: Gesture, segment, prosody* (pp. 68–86). Cambridge: Cambridge University Press.

Bell-Berti, F. (1993). Understanding velic motor control: Studies of segmental context. In M. K. Huffman, & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum. Phonetics and phonology*, vol. 5 (pp. 63–86). San Diego: Academic Press.

Blankenship, B. (1992). What TIMIT can tell us about epenthesis. *UCLA Working Papers in Phonetics* (pp. 17–25) 81.

Bombien, L., Mooshammer, C., Hoole, P., Rathcke, T., & Kühnert, B. (2007). Articulatory strengthening in initial German /kl/ clusters under prosodic variation. In J. Trouvain, & W. J. Barry (Eds.), *Proceedings of the XVIth international congress of phonetic sciences* (pp. 457–460). Saarbrücken: Universität des Saarlandes.

Cambier-Langeveld, T. (1999). The interaction between final lengthening and accentual lengthening: Dutch versus English. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the XIVth international congress of phonetic sciences* (pp. 467–470).

Campbell, L. (2004). *Historical linguistics: An introduction* (2nd ed.). Cambridge, MA: MIT Press.

Cho, T. (2006). Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English. In L. M. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Laboratory phonology 8: Varieties of phonological competence* (pp. 519–548). Berlin/New York: Mouton de Gruyter.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, *3*, 466–485.

Clements, G. N. (1987). Phonological feature representation and the description of intrusive stops. In *Proceedings of the parasession on autosegmental and metrical phonology* (pp. 29–51). Chicago: Chicago Linguistic Society.

Colavin, R. (2007). *Neutralization of epenthetic/underlying [t] contrast in nasal-fricative clusters*. Manuscript, University of California, San Diego. Available online at 〈http://idiom.ucsd.edu/~colavin/epenthesis_comps_1.pdf〉. Downloaded June 12, 2010.

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language*, *49*, 119–132.

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.

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

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

Dembowski, J. (2003). Stop! Don't stop! Epenthesis and assimilation in alveolar clusters. In M.-J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the XVth international congress of phonetic sciences* (pp. 1915–1918). Barcelona: Universitat Autònoma.

Dotevall, H., Lohmander-Agerskov, A., Almquist, S.-Å., & Bake, B. (1998). Aerodynamic assessment of velopharyngeal function during normal speech containing different places of articulation. *Folia Phoniatrica et Logopaedica*, *50*, 53–63.

Erickson, D. (2002). Articulation of extreme formant patterns for emphasized vowels. *Phonetica*, *59*, 134–149.

Fleisch, A. (1925). Der Pneumotachograph: ein Apparat zur beischwindig-keitregistierung der Atem luft. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, *209*, 713.

Fontdevila, J., Pallarès, M. D., & Recasens, D. (1994). The contact index method of electropalatographic data reduction. *Journal of Phonetics*, *22*, 141–154.

Fourakis, M., & Port, R. F. (1986). Stop epenthesis in English. *Journal of Phonetics*, *14*, 197–221.

Hardcastle, W. (1972). The use of electropalatography in phonetic research. *Phonetica*, *25*, 197–215.

Hardcastle, W., Jones, W., Knight, C., Trudgeon, A., & Calder, G. (1989). New developments in electropalatography: A state-of-the-art report. *Clinical Linguistics & Phonetics*, *3*, 1–38.

Harrington, J., Beckman, M. E., Fletcher, J., & Palethorpe, S. (1998). An electropalatographic, kinematic, and acoustic analysis of supralaryngeal correlates of word-level prominence contrasts in English. In *ICSLP-1998*. Paper 0646.

Kilpatrick, C., Shosted, R., & Arvaniti, A. (2007). On the perception of incomplete neutralization. In J. Trouvain, & W. J. Barry (Eds.), *Proceedings of the XVIth international congress of phonetic sciences* (pp. 653–656). Saarbrücken: Universität des Saarlandes.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Lee, S. (1990). The duration and perception of English epenthetic and underlying stops. *Journal of the Acoustical Society of America*, *89*, 1999.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Moon, J. B., & Weinberg, B. (1985). Two simplified methods for estimating velopharyngeal orifice area. *Cleft Palate Journal*, *22*, 1–10.

Mooshammer, C., & Fuchs, S. (2002). Stress distinction in German: Simulating kinematic parameters of tongue tip gestures. *Journal of Phonetics*, *30*, 337–355.

Ohala, J. J. (1974). Experimental historical phonology. In J. M. Anderson, & C. Jones (Eds.), *Historical linguistics* (pp. 353–389). Amsterdam: North-Holland Publishing Company.

Ohala, J. J., & Busà, M. G. (1995). Nasal loss before voiceless fricatives: A perceptually-based sound change. *Rivista di Linguistica*, *7*, 125–144.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer Verlag.

Port, R. F., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, *17*, 257–282.

Story, B., Titze, I. R., & Hoffman, E. A. (1996). Vocal tract area functions from magnetic resonance imaging. *Journal of the Acoustical Society of America*, *100*, 537–554.

Tatham, M. A. (1969). Classifying allophones. *Occasional Papers of the University of Essex Language Centre*, *3*, 14–22.

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

Warner, N., & Weber, A. (2001). Perception of epenthetic stops. *Journal of Phonetics*, *29*, 53–87.

Warren, D. W., Dalston, R. M., & Mayo, R. (1993). Hypernasality in the presence of 'adequate' velopharyngeal closure. *Cleft Palate-Craniofacial Journal*, *30*, 150–154.

Warren, D. W., & Dubois, A. B. (1964). A pressure-flow technique for measuring velopharyngeal orifice area during continuous speech. *Cleft Palate Journal*, *1*, 52–71.

Yoo, I. W., & Blankenship, B. (2003). Duration of epenthetic /t/ in polysyllabic American English words. *Journal of the International Phonetic Association*, *33*, 153–164.