Stop-like modification of the dental fricative /ð/: An acoustic analysis

Sherry Y. Zhao^{a)}

Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 18 March 2009; revised 14 July 2010; accepted 22 July 2010)

This study concentrates on one of the commonly occurring phonetic variations in English: the stop-like modification of the dental fricative /ð/. The variant exhibits a drastic change from the canonical /ð/; the manner of articulation is changed from one that is fricative to one that is stop-like. Furthermore, the place of articulation of stop-like /ð/ has been a point of uncertainty, leading to confusion between stop-like /ð/ and /d/. In this study, acoustic and spectral moment measures were taken from 100 stop-like /ð/ and 102 /d/ tokens produced by 59 male and 23 female speakers in the TIMIT corpus. Data analysis indicated that stop-like /ð/ is significantly different from /d/ in burst amplitude, burst spectrum shape, burst peak frequency, second formant at following-vowel onset, and spectral moments. Moreover, the acoustic differences from /d/ are consistent with those expected for a dental stop-like /ð/. Automatic classification experiments involving these acoustic measures suggested that they are salient in distinguishing stop-like /ð/ from /d/.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3478856]

PACS number(s): 43.70.Fq, 43.70.Bk [AJ]

Pages: 2009–2020

I. INTRODUCTION

Phonetic variations occur widely in speech. Many of the most common variations are still shrouded in mystery in terms of the nature and extent of their change from the canonical forms. Some of these variations pertain directly to the dental fricative /ð/, a phoneme that frequently occurs in spoken English, most often in word-initial positions of function words (Denes, 1963; Maddieson, 1984). In casual speech, /ð/ is frequently modified from its full form (Bell *et al.*, 2003; Jurafsky *et al.*, 1998; Lee, 1990). For example, /ð/ in the function word ‘the’ may be transcribed as [ð], [dh], [d], or [θ]. It may also become nasalized, lateralized, stop-like, or omitted all together (Cao, 2002; Glass, 1988; Manuel, 1995; Manuel and Wyrick, 1999; Roach, 2000; Shockey, 2003; Zhao, 2007).

The stop-like variant of /ð/ is the focus of this study. This modification commonly occurs in word-initial /ð/ that is preceded by silence or a stop consonant (Cao, 2002; Roach, 2000; Shockey, 2003; Zhao, 2007). Occurrences of this stop-like modification have also been observed in various other segmental contexts, such as in word-initial /ð/ that is preceded by fricatives, affricates, liquids or vowels. As shown in Fig. 1, the stop-like /ð/ exhibits a drastic modification from the canonical /ð/; its manner of articulation changes from one that is fricative to one that is stop-like.

Most likely due to the change in production manner, /ð/ is believed to be frequently substituted by /d/ in a number of American-English dialects, such as the New York dialect, Cajun English dialect, and African-American Vernacular English (Dubois and Horvath, 1998; Gordon, 2004; Green, 2002); the substitution also occurs in non-native pronuncia-

tions of English, including those by Italian and Spanish speakers (Goronzy *et al.*, 2001; MacDonald, 1989); and during children’s speech development, often cited as an example of stopping disorder (Bleile, 1995; Marshall and Chiat, 2003). Furthermore, the substitution of /ð/ by /d/ is believed to occur in standard, native American-English, as evidenced by its transcription as an alveolar stop [d] for automatic speech recognition training (Hasegawa-Johnson *et al.*, 2005; Jurafsky *et al.*, 1998).

Despite the seemingly common nature of the /ð/ to /d/ substitution, detailed acoustic comparisons of stop-like /ð/ and /d/, especially those focused on their burst characteristics and adjacent vowel formant movements, are lacking. Such studies may reveal insights to whether stop-like /ð/ differs from /d/ acoustically and the nature of any differences. If salient measures that acoustically separate the two consonants exist, they may be useful for a variety of studies on /ð/, including ones that concentrate on dialects of English, speech acquisition, and speech disorders.

Furthermore, such knowledge would add to our understanding of speech variations, in this case, the nature and extent of the stop-like modification of /ð/. Any measurable acoustic differences between stop-like /ð/ and /d/ may be used to obtain phonemic contrast despite modification. For example, in Stevens’ (2002) model of lexical access from distinctive features, /ð/ and /d/, sharing many of the same articulator-bound features such as [+tongue blade] and [+anterior], differ in the articulator-free feature [continuant]. When /ð/ is produced in a stop-like manner, the distinctive feature [+continuant] cannot be obtained from conventional acoustic measures that differentiate between stop and fricative manners of articulation. However, articulatory gestures that enhance the perceptual saliency of /ð/ may exist and result in measurable acoustic differences from /d/ (Keyser and Stevens, 2006). The existence of such quantitative dif-

^{a)}Author to whom correspondence should be addressed. Electronic mail: syzhao@alum.mit.edu

FIG. 1. Spectrograms of a canonical /ð/ (a) in "buy those" and a stop-like /ð/ (b) in "start those." Utterances are produced by a single male speaker. The canonical /ð/, as highlighted by the brackets in (a), has voicing and weak noises in frequencies above 3 kHz throughout the entire segment. On the other hand, stop-like /ð/ contains a period of no acoustic energy, as highlighted by brackets in (b) suggesting that a complete closure is formed in the oral cavity; it is followed by a sudden acoustic burst, as indicated by the arrow in (b), upon the release of the closure.

ferences may be used to extract the distinctive feature [+continuant] for /ð/ despite a stop-like modification.

In addition to potential contributions toward theories of speech communication, insights from this study may be used to benefit speech applications such as automatic speech recognition (ASR) and speech transcription. Articulatory and/or distinctive feature-based ASR systems may benefit from knowing which acoustic measures to use in order to effectively extract the features of /ð/ despite modification. Conventional ASR systems that extract information from the cepstral domain may benefit from more accurate transcription of /ð/, when the stop-like variant may be identifiably separated from /d/.

Before delving into the experiments and results of this study, an overview of the acoustics of /ð/, along with a review of relevant past studies, are presented in the following sections.

A. Acoustic differences between dental and alveolar stops

/ð/ is a voiced, non-strident, dental fricative. Its production involves vocal fold vibration as well as air turbulence through the vocal tract constriction formed by the tongue between the upper and lower teeth; or the tongue behind the back of the upper teeth. Since there is no complete closure in the vocal tract during its production, /ð/ is continuant; there is continuous acoustic output throughout its production. As previously shown, stop-like /ð/ exhibits a change from a fricative manner of articulation to one that is stop-like. If the dental place of articulation is preserved for stop-like /ð/, then it should exhibit the acoustic characteristics of a dental stop consonant rather than an alveolar stop consonant.

A number of acoustic differences are expected between a dental stop and an alveolar stop. In both cases, the source of excitation is turbulence noise generated near the posterior end of the front cavity. The front cavity is shorter for a dental obstruent than for an alveolar one. Since the natural frequencies of the front cavity are sensitive to small changes in the length dimension, a higher burst spectrum peak location is expected for a dental than for an alveolar. Studies have found

that the main front-cavity resonance of a dental obstruent is around 7 to 8 kHz, while that of an alveolar obstruent is around 4 to 5 kHz (Jongman *et al.*, 2000; Stevens, 1998).

In addition to the difference in burst spectrum peak location, a dental stop is expected to have lower burst amplitude and lower degree of peakedness than an alveolar one. The main factors that contribute to these differences are the length of the front cavity as well as the noise source location and strength. Assuming a simple configuration where the turbulence source is associated with a jet of rapid airflow that exits a constriction, there exists a maximum in spectral prominence corresponding to the front cavity length; experimental data have indicated that when the front cavity length is varied with a turbulence source located 1 cm downstream of the constriction, maximum output is generated at 2.2 cm (front cavity length), which roughly corresponds to the front cavity length of an alveolar obstruent (Stevens, 1989). The amplitude of the spectral prominence decreases when the front cavity is shorter than the optimal length, such as the case for a dental obstruent, due to the increased radiation losses that arise at high frequencies.

Furthermore, the simple model of a noise source associated with rapid airflow exiting a constriction needs to be modified for an alveolar obstruent; its configuration involves an obstacle in the front cavity that creates further turbulence. In the alveolar configuration, the amplitude of the source is increased when rapid airflow through the constriction directly impinges on the back of the teeth (Pastel, 1987; Shadle, 1985). In addition, the obstacle-associated turbulence noise source is located farther from the oral constriction than that for a dental obstruent, where there is no obstacle. A noise source that is farther from the constriction is more effective in exciting the front cavity because it encounters lower acoustic impedance than one located closer to the constriction. Therefore, these acoustic differences are expected to lead to a burst spectrum that has greater high frequency amplitude and is more peaked in shape for the alveolar than for the dental.

Finally, it has been observed that /ð/ has a lower second formant (F2) than /d/ at consonant release in CV contexts for

a number of vowel environments (Cao, 2002; Fowler, 1994; Olive *et al.*, 1993). Furthermore, F2 has been found to be lower for dentals than for alveolars in languages such as Malayalam (Stevens *et al.*, 1986) and O'odham (Dart, 1991), which have phonemically contrastive dental and alveolar stop consonants. One may hypothesize that the difference in F2 is due to a lower, more backed tongue body and a slightly longer back cavity for the dental than for the alveolar; however such a hypothesis must be supported by physiology and articulatory data. In the present study, F2 is expected to be lower for a dental than for an alveolar based on previously mentioned research findings. This acoustic consequence may be associated with distinguishing between laminal (constriction made with tongue blade and thus a more backed tongue body) and apical (constriction made with tongue tip and thus a more fronted tongue body) articulations. Despite the inclusion of F2 in this study, acoustic measures of the release burst spectrum amplitude and shape are also employed to capture differences reflective of the noise source and the obstacle in the front cavity, in addition to the nature of the constriction.

The four different acoustic consequences of the production of dental and alveolar stops, as detailed in this section, stem from different types of acoustic cues. The burst spectrum peak, burst amplitude, and burst spectrum shape describe attributes of the sound that is generated by the turbulent-noise source and the presence of (or lack of) an obstacle in the front cavity of the vocal-tract configuration. The second formant at the onset of the following vowel best represents the influence of the oral-closure release on the transition into the following vowel and may reflect the tongue body shape as well as the back cavity length of the obstruent.

Since the previously mentioned acoustic consequences of dental and alveolar stop consonant production also reflect the vocal tract size and shape of the speaker, they are expected to differ across gender. In particular, gender differences in vocal tract characteristics have been shown to affect both fricative noise spectrum properties and formant frequencies of adjacent vowels (Jongman *et al.*, 2000). Therefore, gender effects need to be carefully considered in analyzing acoustic measures. In the present study, data is reported across all speakers as well as separately for male and female speakers.

In addition to gender effects, dialectal variation may influence any potential acoustic differences between stop-like /ð/ and /d/. For example, acoustic differences between the two consonants may not be expected in dialects with /ð/-to-/d/ substitution, whereas differences may be expected in dialects without the substitution. Therefore, the speaker selection criteria employed in this study include the subject's dialect region to avoid the influence of dialects known to have /ð/-to-/d/ substitution.

B. Past research comparing dental and alveolar stops

The expected differences between dental and alveolar stops presented in this paper are supported by findings of previous studies. Acoustic comparisons of dental and alveo-

lar stop consonants were carried out in studies of several different languages. In Jongman *et al.* (1985) study on the acoustic properties of dental and alveolar stop consonants in Malayalam, the root-mean-square (rms) amplitude of the stop consonant burst normalized to the rms of the adjacent vowel amplitude was found to be salient in distinguishing between the two places of articulation. The normalized amplitude of the alveolar stop burst was consistently higher than that of the dental stop burst. This acoustic attribute was used to distinguish dental stops from alveolar stops in Malayalam with 91.8% accuracy. However, normalized amplitude was not salient in distinguishing between dental stops in Dutch and alveolar stops in American English in the study; 68.2% of the American-English alveolar stops and 63.2% of the Dutch dental stops were classified correctly. Even though American-English alveolar stops were characterized by stronger bursts than Dutch dental stops, speaker-to-speaker variability within each of the two languages seemed to have made accurate classification based on rms burst amplitude difficult. The authors suggested that the greater variability may be explained by the non-contrastive nature of a dental-alveolar distinction in stop consonants of Dutch or English.

Stoel-Gammon *et al.* (1994) found that American-English alveolar /t/ and Swedish dental /t/ differed in voice onset time (VOT), burst intensity and burst spectrum. The VOT values of American-English /t/ were significantly longer than those of Swedish dental /t/. The burst intensity measures were made by taking the following vowel intensity maximum in dB and subtracting the stop-burst intensity maximum in dB; this measure was found to be significantly lower for the American-English alveolar /t/ than for the Swedish dental /t/. In addition, the authors compared the spectral moments, as described by Forrest *et al.* (1988), of American-English /t/ and Swedish /t/ burst spectra. The four spectral moments were used to obtain average energy concentration (mean frequency), spectral shape as indicated by the spread of frequency around the mean (variance used to calculate standard deviation), tilt of the distribution (skewness), and degree of peakedness (kurtosis). The analysis indicated that, compared to Swedish dental /t/, American-English alveolar /t/ was significantly more compact and more peaked.

Sundara (2005) examined the acoustic differences between Canadian-English (CE) alveolar stops /d t/ and Canadian-French (CF) dental stops /d t/. Voice onset time (VOT) was found to be significantly longer for CE alveolar stops than for CF dental stops. Burst intensity was found to be stronger for CE bursts than for CF bursts. Burst spectral shape, measured in terms of mean frequency, standard deviation, and kurtosis, was also different between CE alveolar stops and CF dental dentals. Consistent with the findings of Stoel-Gammon *et al.* (1994), the alveolar stops had higher mean burst frequency, smaller standard deviation, and higher kurtosis than dental stops.

The acoustic differences found between dental and alveolar stops in these studies were consistent with the expected differences based on theoretical derivations. Therefore, if the dental place were preserved for stop-like /ð/, it may be hypothesized that the consonant should have a higher

TABLE I. TIMIT directories (Dir) and selected male (M), female (F) speakers along with number of stop-like /ð/ and /d/ tokens per speaker.

	Dir 2		Dir 3		Dir 7		Dir 8		Total selected dirs	
	M	F	M	F	M	F	M	F	M	F
Total speakers	71	31	79	23	74	26	22	11	246	91
Selected speakers	8	2	13	6	28	8	10	7	59	23
Stop-like /ð/ token per speaker										
1	8	2	12	6	26	6	6	2	52	16
2	–	–	1	–	1	1	3	4	5	5
3	–	–	–	–	1	1	1	1	2	2
Total speakers	8	2	13	6	28	8	10	7	59	23
/d/ token per speaker										
1	6	2	10	5	25	6	6	5	47	18
2	2	–	3	1	2	1	4	1	11	3
3	–	–	–	–	1	1	–	1	1	2
Total speakers	8	2	13	6	28	8	10	7	59	23

burst spectrum peak frequency, lower burst amplitude, less-peaked burst amplitude shape, and lower F2 at the following vowel than /d/. Furthermore, it is hypothesized that these acoustic differences should be expected to exhibit saliency in distinguishing between stop-like /ð/ and /d/. The following sections detail the experiments that test these hypotheses.

II. METHODS

Recorded speech from the TIMIT database was used in this study. The TIMIT database consists of utterances, recorded at 16 kHz sampling rate, from 630 speakers of American English, covering all the major regions of the country (Fisher *et al.*, 1987). For this study, utterances from TIMIT directories 2 (speakers from northern region), 3 (speakers from north midland region), 7 (speakers from western region) and 8 (speakers who moved around the country) were selected for analysis. These directories were selected because they do not include major regional dialects that are known to contain /ð/-to-/d/ substitution, as described in current literature. Some of the regional dialects that are known to contain this substitution are the New York dialect, rural Southern dialects, Cajun English dialect, and Franco-American dialect in New England (Dubois and Horvath, 1998; Gordon, 2004; Ryback-Soucy and Nagy, 2000; Thomas, 2005). Therefore, speakers from New England (directory 1), south midland (directory 4), southern region (directory 5), and New York (directory 6) were excluded from this study.

Table I lists the total number of speakers in the selected TIMIT directories and the actual number of speakers selected for this study. The selection process consisted of examining tokens of /ð/ uttered by each speaker in the directory. Each token was acoustically examined for evidence that suggests the formation and release of a complete oral closure. A stop-like /ð/ segment must contain one of the following series of acoustic events in the speech waveform: 1) a period of silence followed by a release burst; 2) a decreasing voice bar followed by a period of silence before a release burst; 3) or a decreasing voice bar followed by a release burst. It is noted that in this study, stop-like /ð/ tokens preceded by silence or

by a voiceless consonant almost always fit the first criterion, where silence was followed by a release burst; stop-like /ð/ tokens preceded by a vowel or voiced consonant typically fit the second or third criterion involving a decreasing voice bar prior to the closure release.

Speakers who did not produce stop-like /ð/ were excluded from the study. Tokens of /d/ were examined only from speakers who produced stop-like /ð/. Only tokens of /d/ and stop-like /ð/ that were followed by an unstressed vowel, [ə] or [i], were included in the study; this criterion kept the vowel context consistent for all tokens. Therefore, speakers

TABLE II. Segmental context of continuant /ð/, stop-like /ð/, and /d/ tokens.

Context	Continuant /ð/	Stop-like /ð/	/d/
Preceded by silence	16	54	5
Preceding segment across word boundary			
t	0	17	7
d	2	8	1
k	0	2	0
b	0	1	0
v	15	4	1
f	0	0	1
s	1	4	5
z	2	2	8
ð	2	2	0
j	0	2	0
l	7	2	2
r	12	0	18
n	33	0	13
m	8	0	0
ŋ	1	0	5
[ə] or [i]	2	2	19
Other vowels	12	0	0
Preceding segment within same word			
r	0	0	2
l	0	0	2
n	0	0	4
[ə] or [i]	2	0	9

FIG. 2. Burst spectrum peak frequency (F_{hi}) in (a). Ratio of peak amplitude above 2 kHz to that below 2 kHz (A_{hi} -to- A_{lo}) in (b). Spectra with different spectral moment values in (c); the right panel is taken from a /ð/ frication and the left panel is taken from /d/ burst; spectral moments are calculated for regions above 2 kHz in each spectrum.

who produced stop-like /ð/ but did not produce them in the specified vowel context or did not produce any /d/ in similar context were also excluded. The final group of selected speakers all produced at least one stop-like /ð/ token and one /d/ token in the specified vowel context. The resulting 100 stop-like /ð/ and 102 /d/ tokens were obtained from utterances of 23 female and 59 male speakers. The numbers of stop-like /ð/ and /d/ tokens uttered by each speaker are listed in Table I. A few speakers had up to 3 stop-like /ð/ and up to 3 /d/ tokens in the stimulus set; most speakers had 1 token of each consonant. Among the selected speakers, there were far fewer female subjects than male subjects, reflecting the gender imbalance in the TIMIT corpus. In the directories examined, total male-to-female ratio is 2.7 (246 males to 91 females); while the ratio is 2.6 (59 males to 23 females) for selected speakers. Additional female speakers cannot be found without searching in the excluded TIMIT directories.

Table II contains the segmental contexts of stop-like /ð/ and /d/ tokens, as well as that of continuant /ð/ uttered by the selected speakers. All stop-like /ð/ tokens were word-initial; they were most often preceded by silence. On the other hand, continuant /ð/ were most often preceded by sonorants and vowels. Most /d/ tokens were also word-initial; however, the lack of /d/ in unstressed syllables led to the inclusion of non-nasalized /d/ that was preceded by /n/, as well as /d/ that was preceded by either a reduced vowel, a /r/, or a /l/. Unlike stop-like /ð/ tokens, /d/ tokens were most often preceded by a vowel. The contextual differences may be attributed to /ð/ being part of function words such as 'the' and 'this', which are often the first words of TIMIT sentences.

A. Acoustic measures

Acoustic measurements were made using MATLAB. Discrete Fourier Transforms (DFTs) were calculated with a Hamming window of 6.4 ms and a sampling rate of 16 kHz, starting from the beginning of the consonant's release burst to the end of the burst, with 1 ms shifts. An average burst spectrum was taken from the DFTs calculated for each /ð/ and /d/ token. Formant and other vowel spectral measurements were made from the first glottal period at the onset of the vowel following the consonant token. Due to TIMIT's 16 kHz sampling rate, spectral analysis was not conducted beyond 8 kHz.

The first acoustic measure was the high-frequency spectrum peak location (F_{hi}) obtained from the consonant segment. F_{hi} was expected to be higher for the dental place of articulation, which involves a shorter front cavity than for the alveolar place. As illustrated in Fig. 2(a), this variable was measured by first dividing the averaged burst spectrum into eight 1 kHz frequency bands. The burst peak was chosen as the frequency with the highest amplitude in the band with the highest overall energy, above 2 kHz. The approximate dividing point was chosen to be 2 kHz in order to exclude neck-tissue resonance and back cavity resonances excited by vocal fold vibration, or back cavity resonances excited by aspiration if voicing is absent; some low-frequency energies contributed by the front-cavity noise source may also be present due to the coupling of the front and back cavities (Stevens, 1998). Since both stop-like /ð/ and /d/ are voiced, the low-frequency regions of the spectrum were expected to

be similar. Furthermore, energy peak(s) in the low-frequency regions may often be higher than the peak in higher frequencies for voiced obstruents (see Fig. 2). In order to focus on the differences between the two consonants and ensure that the spectrum peak associated with the front-cavity resonance was chosen, analysis of the burst spectrum excluded frequencies below 2 kHz.

The spectrum was calculated using a short-time window of 6.4 ms; frequency components within 300–500 Hz of one another may not be easily distinguished. However, the difference in the burst peak of a dental and an alveolar was expected to be on the order of thousands of Hertz. Thus, this measure should be an adequate approximation of the burst peak frequency differences between the two consonants. Furthermore, an additional measure of the mean frequency in the burst spectrum above 2 kHz, detailed in a subsequent section, was calculated. Any burst peak that was at or higher than 8 kHz could not be obtained due to TIMIT's 16 kHz sampling rate. However, in such cases, the differences in burst spectrum peak and mean frequency of the spectrum, as measured in this study, were still expected to be higher for a dental than for an alveolar.

The second measure, the ratio of the peak amplitude above 2 kHz to that below 2 kHz taken from the averaged release-burst spectrum (A_{hi} -to- A_{lo}), is illustrated in Fig. 2(b). A_{hi} was the amplitude measured at F_{hi} . A_{lo} was the amplitude measured at the frequency with the highest amplitude in the band with the highest overall energy below 2 kHz. Again, the approximate dividing point was chosen to be 2 kHz, in order to separate low-frequency resonances of the back cavity and/or neck-tissues from the high-frequency resonances of the front cavity. Since both stop-like /ð/ and /d/ are voiced, the low-frequency regions of the spectrum were expected to be similar. Thus, they serve as good reference points to normalize the high-frequency regions of the spectrum. This variable was expected to be lower for dental consonants, which should have lower energies in high frequencies than alveolar consonants.

The ratio of total amplitude above 2 kHz in the consonant to that in the following vowel (C-to-V) was also computed. Vowel spectral measurements were made from the first glottal period at the onset of the vowel that follows the consonant token. This variable was expected to be smaller for dentals due to their lower amplitude in high frequencies than for alveolars. Again, the spectral region below 2 kHz was not considered in order to exclude low-frequency energies.

The fourth measure is the second formant frequency (F_2) at following vowel onset. Formant measurements were made from the first glottal period at the onset of the vowel that follows the consonant token. This measure was expected to be lower for the dental than the alveolar, consistent with past research observations and possibly due to a more backed tongue body and longer back cavity for the dental than for the alveolar.

In addition to these acoustic measures, spectral moments were used to obtain measures of energy distribution across frequencies above 2 kHz in the burst spectrum. Spectral moments were first used by Forrest *et al.* (1988) to analyze

voiceless obstruent consonants and had been used to study stop consonants produced by speech-disordered children (Forrest *et al.*, 1990, 1994) among other studies, including ones focused on obstruents (Stoel-Gammon *et al.*, 1994; Flipsen *et al.*, 1999; Jongman *et al.*, 2000; Sundara, 2005). Since both stop-like /ð/ and /d/ are voiced, the spectral region below 2 kHz was not considered in the calculation of spectral moments in order to exclude back cavity resonances and neck-tissue resonances.

The formulas for the spectral moments are listed in Forrest *et al.* (1988). The first spectral moment is the mean frequency or center of gravity of the spectrum. The variance around the mean frequency is the second spectral moment. In the present study, the second moment is used to calculate the standard deviation around the mean. The third moment is the skewness of the spectrum. It is used with the second spectral moment to calculate the dimensionless coefficient of skewness. Negative skewness indicates that the mass of the distribution is concentrated on the right half of the spectrum. Positive skewness indicates that the mass of the distribution is concentrated on the left half of the spectrum. The fourth moment is the kurtosis of the spectrum. It is used with the second moment to calculate the dimensionless coefficient of kurtosis, a measure of the peakedness of the spectrum.

Figure 2(c) contains two spectra with different characteristics and consequently different spectral moment values. The spectrum in the left panel of Fig. 2(c) was obtained from the release burst of a /d/ token; the spectrum in the right panel of Fig. 2(c) was obtained from the voiced frication of a /ð/ token. At regions above 2 kHz, the alveolar release has lower mean frequency and smaller standard deviation around the spectral mean than that of the dental. In addition, the alveolar has higher coefficients of skewness and kurtosis, reflecting a spectral shape that is more concentrated to the left and more peaked than that of the dental. These spectral moment differences are expected when comparing alveolar and dental obstruents.

B. Data analysis

In order to test for significance in the acoustic differences between stop-like /ð/ and /d/, the analysis of variance (ANOVA) was used. Since there were eight different acoustic and spectral measures, a total of eight tests were performed on the collective male-and-female data. Therefore, Bonferroni correction was applied so that the significance level for all eight tests was below an alpha of 0.05; each test employed a significance level of 0.0065. In the analyses of male-only and female-only data, the significance criterion for each of the eight measures was also set to be 0.0065, so that the significance level for all eight tests was below 0.05 for each gender-specific data set. Similarly, in two-way ANOVAs where token group and gender were independent variables, the significance level was again set to 0.0065 for each of the eight tests, so that the collective significance level was below 0.05. Furthermore, the p-values are listed in accompanying data tables for all the significance testing employed.

TABLE III. Acoustic and spectral moment measures for /d/ and stop-like /ð/ across all (A) speakers and separated by male (M) and female (F) speakers.

	/d/	Stop-like /ð/	F ^a	Sig. of F
	Median/mean/stdv	Median/mean/stdv		
F _{hi}	A: 3770 Hz/3730 Hz/1060 Hz	4330 Hz/4480 Hz/1830 Hz	12.47	p < 0.001
	M: 3630 Hz/3520 Hz/870 Hz	4430 Hz/4600 Hz/1720 Hz	22.64	p < 0.001
	F: 4220 Hz/4200 Hz/1320 Hz	3040 Hz/4270 Hz/2050 Hz	0.017	p = 0.89
A _{hi} -to-A _{lo}	A: 1.9 dB/2.5 dB/3.8 dB	0.025 dB/0.2 dB/3.3 dB	19.63	p < 0.001
	M: 1.7 dB/2.0 dB/3.1 dB	-0.04 dB/0.08 dB/3.3 dB	12.72	p < 0.001
	F: 2.3 dB/3.7 dB/5.2 dB	0.39 dB/0.53 dB/3.5 dB	8.0	p = 0.0065
F2 at vowel onset	A: 1830 Hz/1800 Hz/300 Hz	1600 Hz/1670 Hz/330 Hz	8.17	p = 0.0047
	M: 1720 Hz/1730 Hz/310 Hz	1520 Hz/1580 Hz/310 Hz	8.9	p = 0.0033
	F: 1970 Hz/1970 Hz/200 Hz	1840 Hz/1880 Hz/250 Hz	2.24	p = 0.14
C-to-V amplitude	A: 4.1 dB/3.9 dB/4.3 dB	2.9 dB/2.5 dB/6.2 dB	3.48	p = 0.0635
	M: 4.1 dB/3.6 dB/4.2 dB	3.3 dB/2.3 dB/5.9 dB	2.2	p = 0.14
	F: 4.2 dB/4.6 dB/4.7 dB	2.7 dB/2.8 dB/6.7 dB	1.4	p = 0.24
Mean frequency	A: 4180 Hz/4200 Hz/630 Hz	4820 Hz/4860 Hz/480 Hz	69.3	p < 0.001
	M: 4040 Hz/4050 Hz/520 Hz	4890 Hz/4870 Hz/490 Hz	91.8	p < 0.001
	F: 4480 Hz/4580 Hz/710 Hz	4750 Hz/4830 Hz/460 Hz	2.66	p = 0.11
Standard deviation	A: 1340 Hz/1370 Hz/200 Hz	1640 Hz/1640 Hz/170 Hz	113.88	p < 0.001
	M: 1330 Hz/1360 Hz/190 Hz	1620 Hz/1630 Hz/170 Hz	74.1	p < 0.001
	F: 1390 Hz/1390 Hz/200 Hz	1680 Hz/1680 Hz/170 Hz	38.97	p < 0.001
Coeff. of skewness	A: 0.69/0.79/0.67	0.09/0.08/0.38	83.52	p < 0.001
	M: 0.87/0.93/0.58	0.089/0.079/0.38	103.8	p < 0.001
	F: 0.29/0.41/0.74	0.17/0.088/0.38	4.83	p = 0.032
Coeff. of kurtosis	A: 0.19/0.61/1.60	-1.03/-0.91/0.49	83.69	p < 0.001
	M: 0.31/0.75/1.6	-1.01/-0.88/0.49	62.2	p < 0.001
	F: -0.19/0.26/1.4	-1.08/-0.97/0.48	21.09	p < 0.001

^aF(1,200) for all tokens; F(1,139) for tokens by males; F(1,59) for tokens by females.

III. RESULTS

Table III is a summary of the acoustic and spectral moment analysis for 102 /d/ tokens and 100 stop-like /ð/ tokens produced by both male and female speakers and produced by each gender separately. Depending on the measures, the means and standard deviations were rounded to the nearest 10 Hz.

F_{hi} data indicate that the measure is on average lower for /d/ (3730 Hz for all speakers, 3520 Hz for males, 4200 Hz for females) than for stop-like /ð/ (4480 Hz for all, 4600 Hz for males, 4270 Hz for females); this difference is expected since alveolars are produced with a longer front cavity than dentals. The difference in F_{hi} is significant when analyzed across all speakers [F(1,200)=12.47, p < 0.001] and for male speakers only [F(1,139)=22.64, p < 0.001]. However, no significant difference is found for female-only data [F(1,59)=0.017, p = 0.89].

Mean A_{hi}-to-A_{lo} is larger for /d/ (2.5 dB for all speakers, 2.0 dB for males, 3.7 dB for females) than for stop-like /ð/ (0.2 dB for all speakers, 0.08 dB for males, 0.53 dB for females); this is also an expected difference since alveolars should have greater burst amplitude than dentals. The difference in A_{hi}-to-A_{lo} between the two consonants is significant when analyzed across all speakers [F(1,200)=19.63; p < 0.001]. When the data is separated by gender, A_{hi}-to-A_{lo} is still found to be significantly larger for /d/ than for stop-like /ð/ in male-only tokens [F(1,139)=12.72, p < 0.001] as well as in female-only tokens [F(1,59)=8.0, p = 0.0065].

Average F2 at following vowel onset is higher for /d/ (1800 Hz for all speakers, 1730 Hz for males, 1970 Hz for females) than for stop-like /ð/ (1670 Hz for all speakers, 1580 Hz for males, 1880 Hz for females), consistent with F2 differences found between alveolars and dentals in previous studies. In this study, the F2 difference between /d/ and stop-like /ð/ is significant when analyzed across all speakers [F(1,200)=8.17; p = 0.0047] and for male speakers [F(1,139)=8.9, p = 0.0033], but not for female speakers [F(1,59)=2.24, p = 0.14].

The final acoustic measure, C-to-V amplitude, is higher for /d/ (mean of 3.9 dB across all speakers, 3.6 dB for males, 4.6 dB for females) than for stop-like /ð/ (mean of 2.5 dB for all speakers, 2.3 dB for males, 2.8 dB for females). The direction of this difference is expected as alveolars should have greater burst amplitudes than dentals. However, C-to-V amplitude data is not found to be significantly different for the two obstruents, whether it is compared across all speakers [F(1,200)=3.48, p = 0.0635], male speakers only [F(1,139)=2.2, p = 0.14], or female speakers only [F(1,59)=1.4, p = 0.24].

Overall, as expected for a dental obstruent, stop-like /ð/ has on average higher F_{hi}, smaller A_{hi}-to-A_{lo}, lower F2 at following-vowel onset, and smaller C-to-V amplitude than /d/. Analyses of variance revealed statistically significant differences between the two consonants in all measures except C-to-V amplitude.

As for spectral moments, mean frequency is on average lower for /d/ (4200 Hz for all speakers, 4050 Hz for males, 4580 Hz for females) than for stop-like /ð/ (averaging 4860 Hz for all speakers, 4870 Hz for males, 4830 Hz for females). This expected distinction is consistent with the lower F_{hi} observed for /d/ when compared to stop-like /ð/, as both measures reflect the different front cavity lengths involved in alveolar and dental obstruent productions. The difference in mean frequency between the two consonants is found to be statistically significant when compared across all speakers [$F(1,200)=69.3$, $p \leq 0.001$] and when compared across male speakers only [$F(1,139)=91.8$, $p \leq 0.001$]; it is not found to be significant for female speakers [$F(1,59)=2.66$, $p=0.11$].

The standard deviation around the mean is larger for stop-like /ð/ (mean of 1640 Hz for all speakers, 1630 Hz for males, 1680 Hz for females) than for /d/ (mean of 1370 Hz for all speakers, 1360 Hz for males, 1390 Hz for females), an expected difference due to the flatter spectral shape of dentals when compared to that of alveolars. The difference in standard deviation is statistically significant when the comparison is made across all speakers [$F(1,200)=113.88$, $p \leq 0.001$], across male speakers [$F(1,139)=74.1$, $p \leq 0.001$], and across female speakers [$F(1,59)=38.9$, $p \leq 0.001$].

Average coefficient of skewness is higher for /d/ (0.79 for all speakers, 0.93 for males, 0.41 for females) than for stop-like /ð/ (0.08 for all speakers, 0.079 for males, 0.088 for females), suggesting that the former typically have more energy in the left side of the spectrum (2 to 5 kHz) than the latter. The finding is expected as it reflects the longer front cavity and greater excitation associated with alveolar articulation when compared to dental articulation. Skewness dissimilarity between /d/ and stop-like /ð/ is statistically significant when compared across all speakers [$F(1,200)=83.52$, $p \leq 0.001$], as well as across male speakers [$F(1,139)=103.8$, $p \leq 0.001$]; the difference fell short of the specified significance level for female-only data [$F(1,59)=4.83$, $p=0.032$].

The last measure, coefficient of kurtosis, is on average lower for stop-like /ð/ (−0.91 for all speakers, −0.88 for males, −0.97 for females) than that for /d/ (0.61 for all speakers, 0.75 for males, 0.26 for females), suggesting a less peaked shape for the former than the latter. The difference between the two consonants is consistent with theorized differences between dentals and alveolars; it is also statistically significant when compared across all speakers [$F(1,200)=83.69$, $p \leq 0.001$], across male speakers [$F(1,139)=62.2$, $p \leq 0.001$], and across female speakers [$F(1,59)=21.09$, $p \leq 0.001$].

As expected for a dental obstruent, stop-like /ð/ has on average higher mean frequency, larger standard deviation around the mean frequency, smaller coefficient of skewness, and smaller coefficient of kurtosis than /d/. Analyses of variance revealed statistically significant differences between the two consonants in all spectral moment measures.

Two-way analysis of variance was performed for each acoustic and spectral measure, with token group (/d/ and stop-like /ð/) and gender (male and female) as the two inde-

TABLE IV. Main effect of gender on acoustic and spectral measures for all tokens, /d/ tokens only, and stop-like /ð/ tokens only.

	All tokens	/d/	Stop-like /ð/
	F(1, 200)	F(1, 100)	F(1, 98)
F_{hi}	0.59; $p=0.44$	11.22; $p=0.001$	1.01; $p=0.32$
$A_{hi-to-A_{lo}}$	3.88; $p=0.05$	4.24; $p=0.042$	0.39; $p=0.53$
F2 at vowel onset	36.79; $p \leq 0.001$	14.31; $p < 0.001$	22.92; $p \leq 0.001$
C-to-V amplitude	0.78; $p=0.378$	1.05; $p=0.307$	0.13; $p=0.72$
Mean frequency	9.05; $p=0.003$	17.29; $p < 0.001$	0.13; $p=0.71$
Standard deviation	2.37; $p=0.125$	0.45; $p=0.503$	2.54; $p=0.11$
Coeff. of skewness	9.92; $p=0.002$	13.83; $p < 0.001$	0.013; $p=0.91$
Coeff. of kurtosis	2.59; $p=0.109$	1.96; $p=0.16$	0.83; $p=0.36$

pendent variables. Table IV contains the significance of the main effect of gender for each acoustic and spectral measure obtained from both /d/ and stop-like /ð/ tokens. The results indicate that the main effect of gender is significant for F2 at following vowel onset [$F(1,200)=36.79$; $p \leq 0.001$], mean frequency [$F(1,200)=9.05$; $p=0.003$], and skewness [$F(1,200)=9.92$; $p=0.002$]; in other words, these measures, taken from both /d/ and stop-like /ð/, are significantly different between male and female speakers. The main effect of gender is not significant for the rest of the measures when analyzed across all tokens.

The simple main effect of gender for each token group (/d/ or stop-like /ð/) on the acoustic and spectral measures is also contained in Table IV. The simple main effect of gender for /d/ is significant for F_{hi} [$F(1,100)=11.22$; $p=0.001$], F2 at following vowel onset [$F(1,100)=14.31$; $p < 0.001$], mean frequency [$F(1,100)=17.29$; $p < 0.001$], and coefficient of skewness [$F(1,100)=13.83$; $p < 0.001$]; the results indicate that those measures are significantly different between males and females for /d/. The simple main effect of gender for stop-like /ð/ is only significant for F2 at following vowel onset [$F(1,98)=22.92$; $p \leq 0.001$], indicating that it is the only measure that is significantly different between male and female speakers' stop-like /ð/.

The significance of the interaction effects of token group and gender on the acoustic and spectral measures is contained in Table V. There are significant interaction effects of token group and gender on mean frequency [$F(1,200)=12.04$; $p < 0.001$] and skewness [$F(1,200)=10.68$; $p=0.001$]. The interaction effect on F_{hi} [$F(1,200)=6.33$; $p=0.013$] has a p-value that is larger than the significance

TABLE V. Interaction effect of token group (/d/ and stop-like /ð/) and gender on acoustic and spectral measures.

	F(1, 200)	Sig. of F
F_{hi}	6.33	$p=0.013$
$A_{hi-to-A_{lo}}$	1.33	$p=0.25$
F2 at vowel onset	0.60	$p=0.438$
C-to-V amplitude	0.09	$p=0.77$
Mean frequency	12.04	$p < 0.001$
Standard deviation	0.26	$p=0.61$
Coeff. of skewness	10.68	$p=0.001$
Coeff. of kurtosis	1.18	$p=0.28$

level used in the present study but smaller than the conventional threshold of 0.05. There is no significant interaction effect found on the rest of the acoustic and spectral measures.

IV. DISCUSSION

All acoustic and spectral-moment measures except C-to-V amplitude exhibited significant differences between stop-like /ð/ and /d/. The differences were in a manner that was expected if the stop-like /ð/ had dental place of articulation. Compared to /d/, stop-like /ð/ was found to have higher burst spectrum peak and mean spectral frequency, as well as a weaker and flatter spectral burst shape.

All measures exhibited overlap in the data distributions for /d/ and stop-like /ð/. The measure that exhibited the most overlap was C-to-V amplitude; the distribution for /d/ fell within the distribution for stop-like /ð/. The differences in C-to-V amplitude between the two groups were not statistically significant, with a p-value of 0.06. A possible explanation of this smaller than expected difference in C-to-V amplitude may be that most of the /d/ tokens came from unstressed or reduced parts of words such as “departure” and “distress.” Therefore, such /d/ tokens may have lower overall amplitude than what was expected for unreduced cases. This factor may also contribute to the overlap in distributions observed for other measures; in particular those affected by the consonant’s noise-source strength, such as A_{hi} -to- A_{lo} , standard deviation around the mean frequency in the spectrum, coefficient of skewness and coefficient of kurtosis.

F_{hi} was another measure where the distribution for /d/ fell within the distribution for stop-like /ð/. Individual differences in vocal tract characteristics, especially the length, would likely contribute to the variation of F_{hi} . Also, /ð/ was produced with a very short front cavity and weak noise source. Therefore, the amplitude(s) of the front cavity resonance(s) would be small due to the weak excitation and increased radiation losses at high frequencies (Stevens and Keyser, 1989). The resulting relatively flat shape of the stop-like /ð/ burst spectrum would be a factor contributing to the wide distribution of F_{hi} for stop-like /ð/.

The acoustic and spectral moment differences observed in analyses across all speakers were similarly observed in male-only and female-only data. A few measures, F_{hi} , F2, and mean frequency, did not show statistically significant differences for female speakers, despite revealing significant differences for male speakers. The lack of significance in the female-only data may be explained by the small number of tokens produced by female speakers. In addition, the three measures are all influenced by vocal tract resonances; the wide range of individual vocal tract lengths, tongue shapes, and other differences may further contribute to the lack of significance for a relatively small number of tokens.

Consequently, only F_{hi} , F2, and mean frequency displayed varying degrees of distribution overlap when the data is analyzed collectively, containing both male and female speakers, and when it is analyzed separately by gender. More specifically, male-only data displayed the least amount of distribution overlap between stop-like /ð/ and /d/ for each of

the three measures, while female-only data displayed the most overlap. The finding suggests that the distribution overlap observed in the collective male-and-female data for the three measures may be at least partially attributed to the inclusion of both male and female data.

With gender as an independent variable, F2 at following vowel onset was found to be significantly lower for male speakers than for female speakers for both stop-like /ð/ and /d/. This finding is not surprising due to the longer male vocal tract than that of the female. Male and female vocal tract characteristics may also explain the significant gender differences in F_{hi} , mean frequency, and coefficient of skewness observed for /d/ tokens. The shorter front cavity of female speakers may explain their significantly higher F_{hi} and mean frequency than those of male speakers. Consequently, female speakers’ /d/ have significantly lower coefficient of skewness, indicating greater concentration of energy in the higher-frequency half of the spectrum, than those of male speakers. The significantly higher F2, F_{hi} , and mean frequency, as well as significantly lower coefficient of skewness found in female speakers’ /d/ are consistent with previous research on gender differences in the acoustics of fricatives (Jongman *et al.*, 2000).

The lack of significant gender differences in F_{hi} , mean frequency, and coefficient of skewness for stop-like /ð/ may be explained by the consonant being produced with a very short front cavity and weak noise source. As previously mentioned, the weak excitation of the short front cavity, along with the increased radiation losses at high frequencies, lead to a flat and low-amplitude burst spectrum in which resonance-related differences may not be easily captured. Furthermore, the relatively few female speech tokens may contribute to the lack of significance in gender differences for these and other measures examined in this study.

Significant interaction effects between token group and gender were found for mean frequency and coefficient of skewness. Furthermore, the interaction effect would be significant for F_{hi} if the significance threshold were set to 0.05. Each of these three measures display a significant difference between /d/ and stop-like /ð/ for male speakers but not for female speakers, as well as a significant difference between male and female speakers for /d/ but not for stop-like /ð/. Since all three measures are correlates of the front-cavity resonance, the lack of significance for stop-like /ð/ and for female speakers is likely due to the difficulty of capturing resonance-related variation in the flat and low-amplitude noise spectrum of /ð/, which may be further complicated by the relatively small number of tokens produced by female speakers.

Overall, the results of this analysis were consistent with those from previous studies, where dental stops were found to have smaller burst amplitude, smaller coefficients of skewness and coefficients of kurtosis than alveolar stops (Jongman *et al.*, 1985; Stoel-Gammon *et al.*, 1994; Sundara, 2005). However, this study found higher mean frequencies above 2 kHz in the burst spectrum for stop-like /ð/ (median of 4820 Hz for all speakers, 4890 Hz for males, 4750 Hz for females) than for /d/ (median of 4180 Hz for all speakers, 4040 Hz for males, 4480 Hz for females); this finding

seemed to conflict with Sundara's study, where lower mean frequency was found for the dental Canadian-French /d/ [median around 3200 Hz for males; median around 2900 Hz for females, as shown in Sundara (2005), Fig. 4] than for the alveolar Canadian-English /d/ [median around 3800 Hz for males; median around 5000 Hz for females, as shown in Sundara (2005), Fig. 4].

While the male and female mean frequency medians for American-English /d/ were consistent with the medians of the male and female alveolar Canadian-English /d/ analyzed in Sundara's study, the mean frequency medians for stop-like /ð/ were over 1000 Hz higher than the medians for the dental Canadian-French /d/. This discrepancy may be explained by the exclusion of spectral regions below 2 kHz in this study, while Sundara used the entire spectrum to obtain mean frequency. Dental stops were expected to have low energies in high frequencies. Thus, the inclusion of the low-frequency region, which contained energy due to voicing, aspiration if voicing was absent, and/or front-and-back-cavity coupling, would most likely lower the mean frequency of the spectrum.

In order to provide a better comparison with Sundara's data, the entire spectrum was used to recalculate the mean frequency for /d/ and stop-like /ð/. The resulting mean frequency is higher for male speakers' stop-like /ð/ (median of 3070 Hz) than their /d/ (median of 2830 Hz); however the difference did not meet the statistical significance requirement [$F(1,139)=3.233$, $p=0.07$]. As for female speakers, /d/ (median of 3160 Hz) averaged slightly higher than stop-like /ð/ (median of 3140 Hz); the difference was not statistically significant [$F(1,59)=0.29$, $p=0.59$]. While the recalculated mean frequencies of stop-like /ð/ are more consistent with those of dental Canadian-French /d/ obtained by Sundara, the recalculated data for American-English /d/ are much lower than that of alveolar Canadian-French /d/. One possible explanation may be that most of the American-English /d/ tokens came from unstressed or reduced parts of words in order to fulfill the selection criterion of being followed by [ə] or [i]; while the alveolar Canadian-French /d/ tokens were followed by [ɛ], [æ], [o], and [ɔ]. Thus, the American-English /d/ tokens may likely have lower overall amplitudes, where the inclusion of low frequencies would consequently have a greater effect in lowering the mean frequency, than those of Canadian-French tokens. In addition, the lack of significance in the results does not allow for a conclusive comparison of mean frequency in the present study and Sundara's. However, the data seem to suggest that the inclusion of the low-frequency region may not be salient in uncovering differences in mean frequency for American-English /d/ and stop-like /ð/ tokens selected in this study.

While the acoustic measures employed in this study revealed differences between stop-like /ð/ and /d/ in a manner consistent with those expected between dental and alveolar stops, further analysis is needed to substantiate the claim that stop-like /ð/ has the same dental place of articulation as that of canonical /ð/. Such analysis may involve comparisons between the stop-like and the canonical /ð/; careful consideration must be paid to separate acoustic differences due to manner of articulation from those that could be caused by

place of articulation. For example, low-frequency regions for canonical /ð/, which is produced with continuous voicing throughout the segment, would be expected to have higher energy than those for stop-like /ð/. This difference may lead to different spectral characteristics, one taken from the release burst while the other from the frication noise, between the two forms of /ð/ even without any modification of place. Therefore, additional acoustic measures that are not sensitive to manner of articulation differences need to be devised in order to pursue this analysis.

Similar differences in the acoustic and spectral measures examined in this study may be used to distinguish apical and laminal position differences of anterior coronals. Therefore, it may be possible that the differences found in this study were not only due to place of articulation. The possibility of apical and laminal position differences as an alternate explanation of the acoustic differences observed for stop-like /ð/ and /d/ cannot be substantiated in the present study because speakers may possibly produce dentals with both apical and laminal positions (Dart, 1991; Keating, 1991). In order to attribute differences in the previously-mentioned measures solely to place of articulation, consistency in apical or laminal articulations needs to be pursued when comparing stop-like /ð/ and /d/.

Finally, data for all eight acoustic and spectral-moment measures were used in an experiment to gauge their robustness in automatically distinguishing between stop-like /ð/ and /d/. Leave-one-out classification, where one token serves as the unknown test and the rest as training data, was repeated for all 100 stop-like /ð/ and 102 /d/ tokens. The algorithm first normalized the eight acoustic measures and spectral moments by their respective means prior to performing principal component analysis (PCA), which decorrelated the variables of the data set by projecting the data on the eigenvectors of its own covariance matrix (Jolliffe, 1986). This transformation was performed on the collected data across all measures, for all tokens of stop-like /ð/ and /d/. The assignment of unknown tokens was based on the Euclidean distance of each unknown test case to the centroids of the decorrelated stop-like /ð/ and /d/ training data.

Leave-one-out classification using the PCA-transformed data yielded 83% accuracy in classifying the 100 stop-like /ð/ cases and 86% accuracy in classifying the 102 /d/ cases. The results suggested that the eight acoustic and spectral moment measures may be salient in automatically distinguishing stop-like /ð/ from /d/.

V. CONCLUSION

Acoustic measures were obtained for 100 stop-like cases of word-initial /ð/ and for 102 /d/ cases from TIMIT. All cases were followed by an unstressed vowel, [ə] or [i], and extracted from utterances of 23 female and 59 male speakers. Each speaker had at least one /ð/ case and one /d/ case included in this study.

On average, stop-like /ð/ was found to have higher F_{hi} and higher mean frequency in the spectrum above 2 kHz than /d/, both with statistical significance. The differences are expected if stop-like /ð/ has a dental place of articulation

rather than an alveolar one; dentals have higher front cavity resonances because they have a shorter front cavity length than that of alveolars. In addition, stop-like /ð/ was found to have lower F2 at vowel onset than /d/, with statistical significance. This difference is also expected if the modified /ð/ has a dental place, consistent with previous research observations.

Statistically significant differences were also observed in the spectral amplitudes and shapes of the two consonants. On average, stop-like /ð/ had smaller A_{hi} -to- A_{lo} , smaller C-to-V amplitude, larger standard deviation around the mean frequency, smaller coefficient of kurtosis and smaller coefficient of skewness than /d/. These differences are expected if stop-like /ð/ has a dental place of articulation. Dental consonants have lower energies and less-peaked spectral shape in high frequencies than alveolars.

When the tokens were analyzed separately by speaker gender, the expected differences between /d/ and stop-like /ð/ were observed for all eight acoustic and spectral measures. The simple main effect of gender as an independent variable was significant for F2 at following vowel onset, F_{hi} , mean frequency, and coefficient of skewness in /d/, consistent with previous research on gender differences in the acoustics of fricatives (Jongman *et al.*, 2000). However, the simple main effect of gender was only significant for F2 at following vowel onset in stop-like /ð/. The lack of significant gender differences in F_{hi} , mean frequency, and coefficient of skewness for stop-like /ð/ may be explained by the consonant's flat and low-amplitude burst spectrum, where resonance-related differences may not be easily captured. Consequently, significant interaction effects of token group and gender were found for F_{hi} , mean frequency, and coefficient of skewness; the finding may be explained by the lack of significant gender differences in stop-like /ð/ tokens for the three measures, along with the lack of significant differences between /d/ and stop-like /ð/ for the same measures taken from the relatively small number of tokens produced by female speakers.

Overall, the analysis results support the hypothesis that stop-like /ð/ does not exhibit a change in place of articulation to one that is alveolar. Data for most acoustic and spectral measures demonstrate statistically significant differences between stop-like /ð/ and /d/. Consistent with previous research on other modifications of /ð/, this study found acoustic evidence suggesting that the stop-like variant has a dental place of articulation. Furthermore, the eight measures taken from the consonant segment and its following vowel seemed to be robust in automatically distinguishing stop-like /ð/ from /d/. Future extensions of this study may include investigations of human perception of stop-like /ð/ and /d/, place of articulation differences between stop-like /ð/ and canonical /ð/, as well as the effect of apical and laminal articulations on the acoustic differences between stop-like /ð/ and /d/.

In addition, the measures employed in this study may be used to acoustically analyze examples of stop-like /ð/ for evidence of /d/ substitution in English dialects, speech acquisition, speech disorders, as well as for dental-to-alveolar comparisons in languages not limited to English. Also, these measurable acoustic differences may be used to obtain phonemic contrast between the two sounds despite /ð/'s modifi-

cation in manner of articulation. The saliency of these measures suggested that stop-like /ð/ may be an example of enhancement in speech production, where gestures enhance the perceptual saliency of the contrastive feature [+continuant] to provide separability from /d/ despite the modification in manner of articulation.

In conclusion, this study of stop-like modification of /ð/ is an example of the variation in the acoustic and articulatory manifestation of distinctive features that underlie casual speech. Stop-like /ð/ involves a modification of the manner feature, sometimes called an articulator-free feature. Acoustic cues for the implementation of the dental place feature occur in the context of the modified manner of articulation. This type of variation also occurs in utterances such as 'win those', where acoustic cues suggest that the articulator-free feature for the /ð/ segment is [nasal] instead of [continuant] (Manuel, 1995; Stevens, 2002). Cues for the place of articulation appear in the vowel transition and possibly in the nasal murmur as well. Studies like the present one might help to provide some insight into more general principles guiding the types of gestural overlap and acoustic cue modifications that occur in casual speech.

ACKNOWLEDGMENTS

Many thanks to Ken Stevens, Stefanie Shattuck-Hufnagel, Allard Jongman, Michael Kenstowicz, Janet Slifka, and the anonymous reviewers who provided insightful comments on previous versions of this paper. This work is supported by NIH Grant Nos. DC02978 and T32DC00038.

- Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., and Raymond, W. (2003). "Effects of disfluencies, predictability, and utterance position on word form variation in English conversation," *J. Acoust. Soc. Am.* **113**, 1001–1024.
- Bleile, K. M. (1995). *Manual of Articulation and Phonological Disorders: Infancy Through Adulthood* (Thomson Delmar Learning, Clifton Park, NY), pp. 89–135.
- Cao, A. Y. (2002). "Analysis of acoustic cues for identifying the consonant /ð/ in continuous speech," MS thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Dart, S. N. (1991). "Articulatory and acoustic properties of apical and laminal articulations," Ph.D. thesis, University of California, Los Angeles, CA.
- Denes, P. B. (1963). "On the statistics of spoken English," *J. Acoust. Soc. Am.* **35**, 892–904.
- Dubois, S., and Horvath, B. (1998). "Let's tink about dat: Interdental fricatives in Cajun English," *Language Variation and Change* **10**, 245–261.
- Fisher, W., Zue, V., Bernstein, J., and Pallet, D. (1987). "An acoustic-phonetic database," *J. Acoust. Soc. Am.* **81**, S92–S93.
- Flipsen, P., Jr., Shriberg, L., Weismer, G., Karlsson, H., and McSweeney, J. (1999). "Acoustic characteristics of /s/ in adolescents," *J. Speech Lang. Hear. Res.* **42**, 663–77.
- Forrest, K., Weismer, G., Elbert, M., and Dinnsen, D. A. (1994). "Spectral analysis of target-appropriate /t/ and /k/ produced by phonologically disordered and normally articulating children," *Clin. Linguist. Phonetics* **8**, 267–281.
- Forrest, K., Weismer, G., Hodge, M., Dinnsen, D. A., and Elbert, M. (1990). "Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children," *Clin. Linguist. Phonetics* **4**, 327–340.
- Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. M. (1988). "Statistical analysis of word-initial voiceless obstruents: Preliminary data," *J. Acoust. Soc. Am.* **84**, 115–123.
- Fowler, C. A. (1994). "Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation," *Percept. Psychophys.* **55**, 597–610.
- Glass, J. R. (1988). "Finding acoustic regularities in speech: Applications to

- phonetic recognition," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Gordon, M. (2004). "New York, Philadelphia, and other northern cities," in *Handbook of varieties of English: The Americas and Caribbean*, edited by E. Schneider (Mouton de Gruyter, Berlin, Germany), pp. 282–299.
- Goronzy, S., Sahakyan, M., and Wokurek, W. (2001). "Is non-native pronunciation modelling necessary?," in Proceedings of Eurospeech 2001, pp. 309–312.
- Green, L. J. (2002). *African American English: A Linguistic Introduction* (Cambridge University Press, Cambridge, England), pp. 117–172.
- Hasegawa-Johnson, M., Baker, J., Greenberg, S., Kirchhoff, K., Muller, J., Sonmez, K., Borys, S., Chen, K., Juneja, A., Livescu, K., Mohan, S., Coogan, E., and Wang, T. (2005). Landmark-based speech recognition: Report of the 2004 Johns Hopkins summer workshop, Technical Report, Johns Hopkins University, Baltimore, MD.
- Jolliffe, I. T. (1986). *Principal Component Analysis* (Springer, New York), pp. 1–59.
- Jongman, A., Blumstein, S. E., and Lahiri, A. (1985). "Acoustic properties for dental and alveolar stop consonants: A cross-language study," *J. Phonetics* **13**, 235–251.
- Jongman, A., Wayland, R., and Wong, S. (2000). "Acoustic characteristics of English fricatives," *J. Acoust. Soc. Am.* **108**, 1252–1263.
- Jurafsky, D., Bell, A., Fosler-Lussier, E., Girand, C., and Raymond, W. (1998). "Reduction of English function words in switchboard," in Proceedings of the International Conference of Spoken Language Processing, Sydney, Australia.
- Keating, P. A. (1991). "Coronal places of articulation," in *The Special Status of Coronals*, edited by C. Paradis and J.-F. Prunet (Academic, New York), pp. 29–48.
- Keyser, S. J., and Stevens, K. N. (2006). "Enhancement and overlap in the speech chain," *Language* **82**, 33–63.
- Lee, K. (1990). "Context-dependent phonetic hidden Markov models for speaker-independent continuous speech recognition," *IEEE Trans. Acoust., Speech, Signal Process.* **38**, 599–609.
- MacDonald, M. (1989). "The influence of Spanish phonology on the English spoken by United States Hispanics," in *American Spanish Pronunciation: Theoretical and Applied Perspectives*, edited by P. Bjarkman and R. Hammond (Georgetown University Press, Washington, D.C.), pp. 215–236.
- Maddieson, I. (1984). *Patterns of Sounds* (Cambridge University Press, Cambridge, UK), pp. 42–46.
- Manuel, S. Y. (1995). "Speakers nasalize /ð/ after /n/, but listeners still hear /ð/," *J. Phonetics* **23**, 453–476.
- Manuel, S. Y., and Wyrick, G. C. (1999). "Casual speech: A rich source of intriguing puzzles," in 14th International Congress of Phonetic Sciences Conference Proceedings, Vol. 1, pp. 679–682.
- Marshall, C. R., and Chiat, S. (2003). "A foot domain account of prosodically-conditioned substitutions," *Clin. Linguist. Phonetics* **17**, 645–657.
- Olive, J. P., Greenwood, A., and Coleman, J. (1993). *Acoustics of American English Speech: A Dynamic Approach* (Springer-Verlag, New York), pp. 166–175.
- Pastel, L. (1987). "Turbulent noise sources in vocal tract models," MS thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Roach, P. (2000). *English Phonetics and Phonology: A Practical Course*, 3rd ed. (Cambridge University Press, Cambridge, England), pp. 138–142.
- Ryback-Soucy, W., and Nagy, N. (2000). "Exploring the dialect of the Franco-Americans of Manchester, New Hampshire," *Journal of English Linguistics* **28**, 249–264.
- Shadle, C. (1985). "The acoustics of fricative consonants," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Shockey, L. (2003). *Sound Patterns of Spoken English* (Blackwell, Malden, MA), pp. 43–44.
- Stevens, K. N. (1989). "On the quantal nature of speech," *J. Phonetics* **17**, 3–45.
- Stevens, K. N. (1998). *Acoustic Phonetics* (MIT, Cambridge, MA), pp. 398–483.
- Stevens, K. N. (2002). "Toward a model for lexical access based on acoustic landmarks and distinctive features," *J. Acoust. Soc. Am.* **111**, 1872–1892.
- Stevens, K. N., and Keyser, S. J. (1989). "Primary features and their enhancement in consonants," *Language* **65**, 81–106.
- Stevens, K. N., Keyser, S. J., and Kawasaki, H. (1986). "Toward a phonetic and phonological theory of redundant features," in *Invariance and Variability in Speech Processes*, edited by J. S. Perkell and D. H. Klatt (Erlbaum, Hillsdale, NJ), pp. 426–449.
- Stoel-Gammon, C., Williams, K., and Buder, E. (1994). "Cross-language differences in phonological acquisition: Swedish and American /u/," *Phonetica* **51**, 146–158.
- Sundara, M. (2005). "Acoustic-phonetics of coronal stops: A cross-language study of Canadian English and Canadian French," *J. Acoust. Soc. Am.* **118**, 1026–1037.
- Thomas, E. R. (2005). "Rural white southern accents," in *Varieties of English: The Americas/Caribbean*, edited by E. W. Schneider (Mouton de Gruyter, Berlin, Germany), pp. 300–324.
- Zhao, S. Y. (2007). "The stop-like modification of /ð/: A study in the analysis and handling of speech variation," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.