Incomplete Neutralization of /t/ and /d/ in Flapping Environments:
An Analysis of the North Midland Dialect of American English

by

Copyright 2007
Wendy J. Herd

Submitted to the graduate degree program in Linguistics and the
Faculty of the Graduate School of the University of Kansas
In partial fulfillment of the requirements for the degree of
Master's of Arts

Allard Jongman (chairperson)

Joan Sereno (committee member)

Jie Zhang (committee member)

Date Defended _____

The Thesis Committee for Wendy Herd certifies
that this is the approved version of the following thesis:

Incomplete Neutralization of /t/ and /d/ in Flapping Environments:
An Analysis of the North Midland Dialect of American English

Committee:

Allard Jongman (chairperson)

Joan Sereno (committee member)

Jie Zhang (committee member)

Date approved: _____

Acknowledgements

First I would like to thank the two most important people to me, my husband Zac and my daughter Breanna. They have always given me love and support and have been unendingly patient while I try to balance teaching, traveling to school, being a wife and being a mother, although not always in that order. They also made sure I took breaks even when I felt I didn't deserve them. I just hope that I can emulate Breanna's attitude – as she says, "I like to be challenged."

Of course I need to thank my advisor Allard Jongman and my committee members, Joan Sereno and Jie Zhang. In addition to asking me to address challenging questions and always offering useful suggestions, they also helped me with everything from word list issues to statistical analysis stress. Finally, I would like to thank the students from the Empirical Research Seminar for their helpful questions and suggestions.

Abstract

This paper presents an acoustic study of alveolar flaps in American English as produced by twenty speakers of a North Midland dialect spoken in the Kansas area. Vowel duration differences in monosyllabic and disyllabic tokens and consonant duration differences in monosyllabic tokens duplicated previous findings in that vowels preceding /d/ were significantly longer than those preceding /t/ and in that the average duration of /t/ was significantly longer than that of /d/. When analyzing flap frequency, it was discovered that females flapped more often than males and that participants were more likely to flap when they were less aware of the contrast between /t/ and /d/. Contrary to past research, neither word frequency nor morphological complexity affected flap frequency in the present study. Flap frequency was based on a method of distinguishing flapped from unflapped stops on a speaker-by-speaker basis, a more accurate method than that used in previous studies.

1. Introduction

In American English, when a /t/ or /d/ occurs between a stressed and unstressed syllable, it often surfaces as an alveolar flap [ɾ], creating an environment where /t/ and /d/ might be neutralized. According to the previous research of Lehmann and Heffner (1940) and Chen (1970), who examined monosyllabic words contrasting voiced and voiceless consonants in final position, vowel duration is significantly longer preceding voiced segments than voiceless segments and consonant duration of voiceless segments is significantly longer than that of voiced segments. Chen (1970) found mean vowel duration differences of 91 ms and mean consonant duration differences of 52 ms in monosyllabic words of English. This led researchers to expect vowel duration and/or consonant duration differences to aid in the differentiation of /t/ and /d/ when flapped in disyllabic words, avoiding complete neutralization of the two segments.

1.1 Acoustic Analysis of Flaps

Sharf (1962) analyzed the production of an unreported number of monosyllabic words ending in /t/ and /d/, like ‘cat’ and ‘cad’, and disyllabic words with medial /t/ and /d/ in flapping environments, ‘catty’ and ‘caddy’ for example. He found that vowels were 91 ms longer on average before word-final /d/ than /t/ in monosyllabic words, and there was a 9 ms difference in disyllabic words with flaps. In contrast to Chen’s (1970) findings, where the consonant durations of voiceless segments were longer than their voiced counterparts, Sharf (1962) found that the

average closure duration of flapped /d/ (30 ms) was only 4 ms longer than the average closure duration of flapped /t/ (26ms).

In a similar acoustic study, Fox and Terbeek (1977) examined the production of 10 disyllabic pairs like ‘putting’ and ‘pudding’ produced by 21 native speakers of English (10 male and 11 female). Like Sharf (1962), Fox and Terbeek (1977) found that although /t/ and /d/ appear to neutralize when flapped, vowels preceding /d/ flaps were significantly longer than those preceding /t/ flaps and the closure durations of /d/ flaps and /t/ flaps were not significantly different, but the sizes of the differences were not reported. In order to account for variation in speech rate, Fox and Terbeek (1977) also measured and compared the first vowel/second vowel ratios of /t/ flap words and /d/ flap words, finding that the ratios were significantly higher for /d/ flap words, but again the size of the effect was not reported.

Zue and Laferriere (1979) reported similar vowel duration and closure duration results when they analyzed the production of 250 polysyllabic words read by six native English speakers (three male and three female). In addition to analyzing environments where /t/ and /d/ were produced as flaps, Zue and Laferriere (1979) also looked at /t/ and /d/ in five other environments, so the 250 tokens were not limited to flapping environments. In their analysis of flapped /t/ and /d/, Zue and Laferriere (1979) documented that vowels were 9 ms longer preceding /d/ flaps than /t/ flaps, a statistically significant difference that replicated Sharf’s (1962) results. Also like Sharf (1962) and Fox and Terbeek (1977), Zue and Laferriere (1979) noted only a 1

ms difference between /d/ flap closure duration (27 ms) and /t/ flap closure duration (26 ms), a statistically insignificant difference.

More recently, Patterson and Connine (2001) searched the SWITCHBOARD corpus for minimal pairs like ‘putting’ and ‘pudding’ in which /t/ and /d/ are flapped, finding 11 minimal pairs, 60 tokens with /t/ flaps and 55 with /d/ flaps. Like previous researchers, Patterson and Connine (2001) found that average length of vowels preceding /d/ flaps was 16 ms longer than that of vowels preceding /t/ flaps, a statistically significant difference. In addition to replicating the results of past research, they noted that flapping occurred significantly less often when the tokens were low frequency words and when the tokens were morphologically complex, like ‘beating’.

While Sharf (1962), Fox and Terbeek (1977), Zue and Laferriere (1979), and Patterson and Connine (2001) documented vowel length differences that acoustically distinguish an underlying /t/ flap in a word like ‘writer’ [ˈɹaɪt̬ə] from ‘rider’ [ˈɹaɪd̬ə] with an underlying /d/, no attempt was made to look at a specific dialect of American English. The one speaker in the Sharf (1962) study was from Detroit, but Fox and Terbeek (1977) and Zue and Laferriere (1979) refer to their participants only as speakers of an American English dialect. In the Patterson and Connine (2001) study, the minimal pairs were drawn from the SWITCHBOARD corpus without restriction to one regional dialect, and with individual tokens within minimal pairs produced by different speakers, where vowel duration differences could be due in part to speaker

variation. Additionally, these studies included no perceptual tests to evaluate how salient the vowel differences found are to listeners.

1.2 Perceptual Studies of Flaps

One of the earliest perceptual studies was conducted by Sharf (1960). After recording two speakers (one male and one female) who read a pseudorandomized list of 15 minimal pairs with underlying /t/ and /d/ in flapping environments, Sharf (1960) played the recordings for 12 native speakers of American English. The listening participants were instructed to check the word they thought they heard on a numbered list of target words with minimal pair counterparts. As the female speaking participant did not flap /t/ in her productions, listeners were able to correctly perceive her productions 86.5% of the time. Alternately, listeners were only able to correctly perceive the male speaker's productions, the productions which truly contained flaps, 64% of the time. While this is above chance, Sharf (1960) did not report a statistical analysis of his results. Although Sharf (1960) does not describe how it was determined if a speaker were producing an alveolar stop or a flap, he noted that the female speaker "used a [t] sound" (107).

Later, Malécot and Lloyd (1968) administered a perception test in which they randomized an unreported number of minimal pairs with medial /t/ and /d/ produced by five native speakers of an Eastern American dialect. Fifty native English-speaking participants took the test which required them to circle the correct spelling of the word they heard, so participants would hear ['kiri] and then circle either 'kitty' or 'kiddie', a similar format to that used by Sharf (1960). The listeners performed near

chance levels, averaging 56.6%, and researchers found that listeners had a general /d/ bias, probably due to the voiced quality of flaps. After noticing the /d/ bias in listeners' responses, Malécot and Lloyd (1968) visually examined the spectrograms of the minimal pairs, looking for a correlation between vowel length and percent of /d/-responses by listeners. Although no vowel durations were reported, they did report a strong positive statistical correlation between the length of monophthong vowels and /d/-judgments; however, no such correlation was found for diphthongs.

These two early perceptual studies both fall short, Sharf (1960) in its use of only two speakers, one of whom did not produce /t/ as a flap and both studies in their lack of definitive results as to listeners' abilities to distinguish /d/ and /t/ when flapped. While the findings of Malécot and Lloyd (1968) indicate that listeners may have used monophthong vowel length as a cue for their judgments, they did not confirm that the speakers made significant or consistent vowel length differences when producing words with /d/ and /t/ medially. The last group of studies discussed combines both acoustic analyses and perception tests to investigate whether /t/ and /d/ are neutralized when flapped.

1.3 Acoustic and Perceptual Studies of Flaps

Fisher and Hirsh (1976) completed an early acoustic and perceptual study investigating whether /t/ and /d/ were neutralized when flapping. They recorded six different native speakers of American English reading 18 disyllabic minimal pairs containing medial /t/ and /d/ and 18 monosyllabic pairs containing word-final /t/ and /d/. During the acoustic portion of the study, the vowel length differences of only

three of the speakers were analyzed because they “were the most natural and reliable speakers” (192). Based on these three speakers, Fisher and Hirsch (1976) found that vowel length differences before /t/ and /d/ were reduced when going from monosyllabic to disyllabic words; however, vowels were still longer preceding /d/ flaps than preceding /t/ flaps. An analysis of the closure duration of flapped /t/ and /d/ for all six speakers found that flapped /d/ was actually longer in duration than flapped /t/, but not significantly so, replicating the results of Sharf (1962) and Zue and Laferriere (1979). During the perception portion of the study, the recordings were played for six phonetically trained judges who transcribed the medial /t/ and /d/ as [t^h], [t], [d], [r], voiced [r], or voiceless [r]. Based on the transcriptions of these six judges, who could reliably distinguish between voiceless and voiced flaps, Fisher and Hirsch (1976) concluded that /t/ and /d/ are not neutralized in flapping environments. Unfortunately, the acoustic results of this study are based on only a small number of speakers, with the vowel length data based on only half of those participants, and the perceptual data is based on the judgments of trained phoneticians, not telling us whether the difference between flapped /t/ and flapped /d/ is salient to naïve speakers of American English.

Charles-Luce (1997) conducted an acoustic and perceptual study that examined how the manipulation of listener-present/listener-absent and biased passage/neutral passage conditions affected the /t/ and /d/ contrast in flapping environments. Ten speaking participants (5 male, 5 female) read 20 randomized passages containing five minimal pairs contrasting /t/ and /d/ in a flapping

environment. The five minimal pairs, or ten words, were read once in biased passages, narratives in which the target word was predicted, and once in neutral passages, unrelated sentences that did not predict the target. The ten participants were divided into two conditions: listener-present, where the participant read the 20 passages with another person present, and listener-absent, where the participant read the 20 passages without a listener present. Based on the measurements of closure duration, first vowel duration, and overall word duration, Charles-Luce (1997) first found, like past researchers, that closure duration was never a significant indicator of voicing. Vowel duration differences, on the other hand, significantly predicted /t/ and /d/ flaps only in the biasing passages and not in the neutral context, indicating that speakers maintained the voicing difference when the target word was predictable. Similarly, the voicing difference between /t/ and /d/ was maintained in the listener-present condition through overall word duration, with the word duration being longer for /d/ flap words and shorter for /t/ flaps. Charles-Luce (1997) concedes that no significant vowel duration or word duration differences were found independent of the two conditions: biased/neutral passage and listener-present/listener-absent. The perceptual study yielded more surprising results. Twenty-four native speakers of American English listened to 192 disyllabic words in isolation read by the 10 speakers in the two conditions, listener-present and listener absent. Listeners were instructed to choose the word they heard from a response booklet, so they would hear 'writer' and would choose 'rider' or 'writer' from the response booklet. The listeners responded correctly 81% of the time, 80% for listener-absent and 82% for listener-

present, an astoundingly high percentage correct when compared to the results of Sharf (1960) and Malécot and Lloyd (1968). Although a main effect of listener-condition (listener present or absent) was reported, passage type (biased or neutral) was not significant. While Charles-Luce does not comment on the disparity between the results of her perceptual study and that of previous researchers, the varying results indicate a need for further research.

In similar studies of semantic predictability, Charles-Luce and Dressler (1999) and Charles-Luce, Dressler and Ragonese (1999) returned to the question of how semantic predictability, but not listener condition, affected the production and perception of /t/ and /d/ when flapped. The studies found that children near the age of seven and older adults tended to use closure duration to indicate voicing, apparently hyperarticulating to the point that they were no longer producing flaps. Alternately, both younger adults and children near the ages of nine and twelve produced vowel duration differences before /t/ and /d/ when flapped in biased passages only, reproducing the results of Charles-Luce (1997). During the perceptual portion of the studies, the hyperarticulated productions of the children near age seven and the older adults were unsurprisingly more accurate than that of the other children and the young adults. More interesting, the percentage of tokens perceived correctly ranged from 56% to 73% across all groups, considerably less than that reported in Charles-Luce (1997).

The acoustic and perceptual studies of flaps conducted over the past forty years leave questions that have yet to be answered. While acoustic studies have

consistently shown that the closure duration of /t/ and /d/ when flapped is not an accurate cue to voicing, studies vary over the description of preceding vowel duration as a cue to voicing. Earlier studies documented a 9 ms – 16 ms difference in vowel duration, where vowels were longer preceding /d/ flaps, but the more recent Charles-Luce (1997) study found that neither vowel duration nor word duration were statistically significant independent cues. Furthermore, researchers have yet to define what constitutes a flap, using various arbitrary measures, like 50 ms, 10 – 40 ms, or half the duration of an average [t] or [d]. When a description is given as to how a /t/ or /d/ is determined to be flapped, it is usually in the form of a vague reference to auditory judgments by the researcher(s) and visual analysis of phonetic measures, presumably waveforms and spectrograms. Similarly, the perceptual studies have varied dramatically. While some studies, like Malécot and Lloyd (1968), have reported a percentage of tokens perceived correctly very near chance levels, others, like Charles-Luce (1997), have documented considerably higher levels of accuracy.

In addition to the question of whether /t/ and /d/ are acoustically or phonetically neutralized in flapping, the studies have not considered that regional dialect differences may have lead to the varying results. No one has set out to study the production and perception of vowel and closure duration differences in /t/ and /d/ flaps within one regional dialect. As a portion of their study, Patterson and Connine (2001) compared the frequency of flapping between the speakers of the Northern dialect and the North Midland dialect of American English. Although they and Byrd (1994) predicted that no variation in flapping would occur across dialects, both

studies reported differences in flap frequency. Byrd (1994) found that American English speakers from the North and the Northeast and speakers from the North Midland produced no significant variation in flapping the word ‘water’, but the speakers from the North and the Northeast were significantly less likely to flap across morpheme-boundaries in the phrase ‘suit in’ than their North Midland counterparts. Byrd (1994) attributed this to a difference between word-medial and word-final flapping, but the similar flapping frequency in ‘water’ across dialect regions may have had more to do with the word frequency of ‘water’, which has a Kučera-Francis written word frequency of 442 (Francis and Kučera 1982). Like Byrd (1994), Patterson and Connine (2001) reported that speakers of the Northern dialect of American English produced flaps less often, with 25.2% of tokens being unflapped, than speakers of the North Midland dialect, who produced only 18.4% of flappable tokens without flaps. Unfortunately, Patterson and Connine (2001) did not report whether or not the difference in flap frequency was significant. Since past research indicates that flaps can differ in frequency across dialect regions, it is also possible that the vowel durations preceding those flaps and the durations of the flaps themselves differ.

This paper attempts to use acoustic measurements in order to investigate whether /t/ and /d/ have neutralized in flapping environments for speakers of the North Midland dialect of American English. Based on previous research, I expect the average vowel duration differences before flapped /t/ and /d/, if maintained by this group of speakers, to be near the range of 7 – 16 ms, much shorter than vowel length

differences preceding /t/ and /d/ found in monosyllabic words. Like previous researchers, I expect to find no difference in the duration of the flaps based on whether they were underlying /t/ or /d/.

2. Acoustic Study

2.1 Method

Participants

Thirty native American English speakers from the University of Kansas in Lawrence participated in the acoustic portion of this study. Of these speakers, five males and three females were excluded from analyses, because they reported living more than 50 miles outside of Kansas on the dialect survey, which was administered to all participants. Two other female speakers, the last two recorded, were also excluded prior to analyzing the data in order to balance gender within the participant groups. Analyses discussed in this study are based on twenty speakers, ten females and ten males, who were born in or within ten miles of Kansas and who had not lived outside of the area for more than one year. All participants were recorded reading three different lists of words and nonwords containing word-final and word-medial /t/ and /d/.

Stimuli

The lists contained 28 disyllabic minimal pairs contrasting /t/ and /d/ in flapping environments, half of these with diphthongs as the first vowel and half with monophthongs. Based on Sharf (1960) and Malécot and Lloyd (1968), diphthongs may not show the same vowel duration differences as monophthongs. Additionally,

the lists included 26 monosyllabic minimal pairs contrasting /t/ and /d/ word-finally.

The monosyllabic pairs were counterparts to the disyllabic minimal pairs, so the disyllabic pairs ‘tidal’-‘title’ and ‘tidy’-‘tighty’ had the corresponding pair ‘tide’-‘tight’. Because two of the monosyllabic pairs, ‘tide’-‘tight’ and ‘wide’-‘white’, matched up with two disyllabic pairs each, there were 28 disyllabic minimal pairs but only 26 monosyllabic as seen in Table 1 below. When no monosyllabic counterparts

Table 1. Lexical monosyllabic and disyllabic minimal pairs.

Vowel	Word Final /d/	Word Final /t/	Intervocalic /d/	Intervocalic /t/
aɪ	ride	write	rider	writer
	tide	tight	tidal	title
	wide	white	widest	whitest
	bide	bite	biding	biting
	side	sight	sided	sighted
	tide	tight	tidy	tighty
	wide	white	wide-ish	white-ish
eɪ	grade	grate	grader	grater
	raid	rate	raiding	rating
	fade	fate	faded	fated
aʊ	cowed	pout	powder	pouter
	cloud	clout	clouded	clouted
oʊ	bode	boat	boded	boated
	code	coat	coding	coating
i	Swede	sweet	Swedish	sweetish
	lead	cleat	leader	liter
ɪ	kid	kit	kiddie	kitty
	bid	bit	bidder	bitter
ɛ	bed	pet	pedal	petal
	wed	wet	wedding	wetting
æ	mad	mat	madder	matter
	pad	pat	padding	patting
ʌ	bud	but	budding	butting
	thud	shut	shudder	shutter
u	tude	toot	Tudor	tutor
	rude	root	Rudy	rooty
ɔ	odd	ought	odder	otter
	plod	plot	plodding	plotting

were available for disyllabic minimal pairs, for example in the minimal pair ‘pouter’ - ‘powder’, a near minimal pair was used, ‘pout’ - ‘cowed’. The monosyllabic minimal pairs allowed the researcher to measure an average [t] and [d] duration for each speaker, enabling her to better define what would be considered a flap for each individual speaker and allowing verification that speakers maintained vowel duration differences in monosyllabic words. Additionally, 28 disyllabic words that do not

Table 2. Lexical items with nonword minimal pairs with the lexical items in bold.

Vowel	Intervocalic /d/	Intervocalic /t/
aɪ	glider	glighter
	guided	gighted
	vaidle	vital
	fider	fighter
eɪ	jaded	jaited
	lady	laety
	laeder	later
	fadle	fatal
aʊ	chowder	chouter
	crowded	crouted
	owded	outed
	spowded	spouting
oʊ	odor	oater
	voder	voter
i	greedy	greety
	sweedy	sweetie
ɪ	kidder	kitter
	glidder	glitter
ɛ	shedding	shetting
	swedder	sweater
æ	daddy	datty
	baddy	batty
ʌ	puddle	puttle
	shuddle	shuttle
u	brooding	bruting
	shooder	shooter
ɔ	modest	mottest
	hoddest	hottest

exist in minimal word pairs were used along with 28 nonword counterparts, for example ‘shedding’ and ‘shetting’. Of the 28 disyllabic nonminimal pairs shown in Table 2 above, the medial consonant was /d/ in half of the lexical items and /t/ in the other half. Analysis of the listeners’ judgments of this list during the perceptual study should shed light on the interaction of /d/-bias, word-frequency, and lexical status.

In addition to balancing monophthongs and diphthongs in the word list, vowel quality, word frequency, and morphological complexity were also considered. An attempt was made to include a variety of vowels and suffixes. Using the Kučera-Francis (1982) written word frequencies for the disyllable lexical items, a paired samples t-test conducted comparing words with underlying /t/ ($\bar{x} = 33$, s.d. = 85.5) and /d/ ($\bar{x} = 15.24$, s.d. = 34.2) found that the two lists did not significantly differ in word frequency [$t(41) = -1.214$, $p = .232$]. Likewise, when another paired samples t-test was performed comparing the number of morphemes in words containing /t/ ($\bar{x} = 1.64$, s.d. = .49) and /d/ ($\bar{x} = 1.74$, s.d. = .45), the number of morphemes in the two lists did not differ significantly [$t(41) = 1.071$, $p = .290$].

Procedure

In the first context, the speakers read a list consisting of the 164 target words and nonwords pseudorandomized with 51 fillers. During context two, speakers read the same list with the words inserted into the carrier sentence, ‘Please say _____ again.’ For the third context, speakers read a paired flap list, in which the 164 tokens were listed as minimal pairs and were pseudorandomized with 20 fillers. Half of the

participants, five males and five females, read the list with /t/ tokens first and half, five males and five females, with /d/ tokens first. Speaking participants were instructed to speak as naturally as possible. The lists were constructed in such a way that speakers would be unaware of the goal of the study in context one's word list and context two's carrier list, resulting in natural speech, but the speakers would be focused on the difference between minimal pairs in the paired flap list, possibly resulting in more salient differences between /t/ and /d/.

Speakers were recorded in the anechoic chamber at the University of Kansas in Lawrence. One speaker was recorded on DAT tape and later digitized using PRAAT software at a sampling rate of 22 kHz with a low-pass filter setting of 10 kHz. The other nineteen speakers were recorded onto a flash disc using a Marantz Portable Solid State Recorder at a sampling rate of 22 kHz.

2.2 Measurements

Four acoustic measurements were taken for each monosyllabic word or nonword using PRAAT software: word length, first vowel duration, a ratio of vowel duration to word length, and /t/ or /d/ duration. In addition to these four measurements, a first vowel to second vowel ratio was also taken for disyllabic words and nonwords. Figure 1 below illustrates where these measurements were made.

For tokens read in isolation or in pairs, word length was measured from the onset to the offset of the visible waveform and spectrogram, and for tokens read in the carrier sentence, word length was measured between changes in the cyclic nature of the waveform and changes in the spectrogram at word boundaries. The first vowel

for all words and nonwords and the second vowel for disyllabic words and nonwords were measured from the onset of the first formant to the offset of the second as seen in spectrograms. The /t/ or /d/ duration for monosyllabic tokens was measured from the offset of the first vowel to the end of the visible waveform and the end of the stimuli in the spectrogram when items were read in isolation or in pairs. When read in the carrier sentence, the end of the consonant duration was measured as the change in the cyclic nature of the waveform and a change in the spectrogram at the word

Figure 1. Measurements of word length and vowel duration in the word ‘later’.

boundary. For disyllabic items, the /t/ or /d/ duration was measured from the offset of the first vowel to the onset of second.

3. Results

3.1 Monosyllables

Paired samples t-tests were conducted comparing the word-final consonant duration, vowel duration, word duration, and vowel to word length ratio of the monosyllabic /t/ and /d/ pairs. As found in previous research and presented in Figure 2 below, the mean duration of voiceless /t/ was 82 ms longer than the mean duration of voiced /d/ [$t(814) = 35.51, p < .0001$]. Similarly, the mean vowel duration preceding /d/ was 92 ms longer than the mean vowel duration preceding /t/ [$t(814) = 63.34, p < .0001$]. Likewise the mean word length of monosyllabic words ending in

Figure 2. Vowel duration preceding word-final /t/ and /d/ and consonant duration of word-final /t/ and /d/.

/d/ measured 13 ms longer than words ending in /t/ as illustrated in Figure 3 below [$t(814) = 4.009, p < .0001$]. The word final /d/ tokens are 13 ms longer than the /t/ tokens because the difference in vowel duration, where /d/-final words have average vowel durations 92 ms longer than /t/-final words, is 10 ms longer than the difference

in consonant duration, where /t/-final words have average final consonant durations 82 ms longer than /d/-final words. Similarly and also due to the longer vowel

Figure 3. Mean word duration in monosyllabic words ending in /t/ and /d/.

duration preceding word-final /d/, the mean vowel to word length ratio of /d/-final words was 0.17 greater than that of /t/-final words [$t(814) = 60.556, p < .0001$] as seen in Figure 4 below.

Figure 4. Mean vowel duration to word length ratio of monosyllabic words ending in /t/ and /d/.

The mean vowel duration and consonant duration differences reported above demonstrate that these twenty native English speakers maintain vowel and consonant duration differences in monosyllabic words. Furthermore, the mean consonant duration of each speaker's alveolar stops establishes a baseline against which one can compare the flap durations on a speaker-by-speaker basis. As seen in Tables 3 and 4 speaker-by-speaker variations in consonant duration and the difference in duration of /t/ and /d/ as well as variations in vowel duration and the difference in duration preceding /t/ and /d/ exist. Likewise, one would expect a variation in flapped consonant duration to exist, reinforcing the need for a more objective definition of flapped /t/ and /d/ than an arbitrary cut-off point like 50 ms for all speakers.

Table 3. Speaker variation in consonant duration for monosyllables.

Speaker #	/t/ Duration (ms)	s.d.	Minimum	Maximum	/d/ Duration (ms)	s.d.	Minimum	Maximum	Difference (ms) (/t/ - /d/)
1	214	61.6	83	385	153	34.7	89	280	61
2	195	52.4	102	377	136	28.6	83	218	59
3	236	46.2	151	359	167	39.7	77	300	69
4	289	47.0	125	387	157	34.6	70	252	132
5	223	70.1	142	391	136	28.1	82	185	87
6	140	27.6	78.6	195	107	21.4	62	147	33
7	217	25.6	162	315	124	20.8	91	182	93
8	230	42.5	166	286	114	20.1	93	146	115
9	204	37.6	98	319	121	19.0	59	151	83
10	165	26.5	142	200	107	22.8	91	147	58
11	232	31.3	180	307	151	23.7	105	218	81
12	228	52.5	103	373	139	28.2	91	220	89
13	219	42.2	115	316	126	23.3	73	178	93
14	181	75.4	113	420	140	34.1	79	265	41
15	234	57.5	107	349	148	27.1	74	224	86
15	171	49.0	78	309	108	25.5	71	182	63
17	158	68.6	103	303	100	21.1	70	137	58
18	232	40.1	162	340	113	17.1	78	162	119
19	168	39.8	118	205	125	33.7	102	174	43
20	184	55.9	96	337	126	36.0	92	355	58

Table 4. Speaker variation in vowel duration for monosyllable.

Speaker #	Vowel Duration preceding /t/ (ms)	s.d.	Minimum	Maximum	Vowel Duration preceding /d/ (ms)	s.d.	Minimum	Maximum	Difference (ms) (/d' - /t/)
1	180	45.9	62	261	259	50.8	79	347	79
2	154	38.1	93	236	240	54.0	139	354	86
3	161	39.7	84	261	254	54.8	161	380	93
4	191	41.4	165	284	262	51.7	105	284	71
5	189	34.8	115	247	254	56.0	147	360	65
6	148	31.2	75	224	245	43.3	164	323	97
7	172	28.2	108	226	259	42.4	168	330	87
8	174	32.9	120	207	280	56.9	194	347	106
9	166	29.4	102	249	305	52.4	184	402	139
10	158	39.3	95	200	225	73.4	111	310	67
11	165	31.4	106	238	256	44.6	150	362	91
12	165	42.3	61	242	258	55.0	89	361	93
13	160	35.5	90	233	245	50.5	137	343	85
14	195	42.8	113	267	270	45.9	168	352	75
15	167	40.8	66	264	286	44.7	197	387	119
16	197	34.1	151	251	254	32.9	195	290	57
17	160	30.9	120	200	211	53.7	157	289	51
18	183	35.4	111	266	304	49.2	217	412	121
19	184	35.8	142	230	268	25.5	239	297	84
20	159	27.5	93.5	219	235	41.3	155	342	76

3.2 Defining Flaps

As mentioned above, previous studies used one value, usually a number between 30 and 50 ms, as a boundary between alveolar flaps and stops for all speakers, a method which fails to take speaker variability into account. Instead of using an arbitrary number to define flaps across all speakers, each speaker's productions were analyzed separately in this study. First, the consonant durations of all of a speaker's monosyllabic and disyllabic tokens (156 monosyllabic words and 336 disyllabic words and nonwords) were graphed, as shown below in Figure 5, resulting in a bimodal distribution. Next, based on a gap in the distribution between

Figure 5. Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens for Speaker 3.

flapped and fully articulated stops, a value distinguishing the two was assigned to each speaker. As can be seen in Figure 5, the distribution for Speaker 3, a clear boundary falls between 52 ms and 62 ms, so all tokens below this boundary were considered flapped and all above were considered unflapped for this speaker. For all speakers, there was at least one empty bin separating flapped from unflapped stops as can be seen in Appendix A. Table 5 below lists the value at which flaps were defined for each of the twenty speakers. The average flap cut-off value across all twenty speakers was 56.25 ms, with a range of 43 – 69 ms.

Table 5. Value at which each speaker's /t/ and /d/ productions were considered flaps.

Speaker Number	Flap cut-off value (ms)
1	57
2	61
3	52
4	61
5	61
6	58
7	53
8	43
9	57
10	43
11	60
12	60
13	53
14	60
15	69
16	47
17	57
18	59
19	55
20	59

In order to verify that a clear boundary fell between flapped and unflapped alveolar stops in both monosyllables and disyllables and for both voiced and voiceless segments, separate distributions were generated. For example, Figure 6a below represents the distribution of all word-final /t/ and /d/ tokens produced by Speaker 3 in monosyllables and clearly illustrates a gap between the flapped and unflapped tokens between the 50 ms and 75 ms labels. The flapped monosyllabic tokens, which were produced in a flapping environment due to the carrier sentence ‘please say _____ again’, are shown enlarged in Figure 6b in order to confirm the regular distribution of /t/ and /d/ tokens labeled as flaps. Separate distributions of all monosyllabic /t/ tokens and /d/ tokens are shown in Figures 6c and 6d respectively. Both distributions demonstrate that the value 52 ms acts as an effective boundary between flapped and unflapped alveolar stops.

Figure 6. Distributions of all /t/ and /d/ tokens (a), flapped /t/ and /d/ tokens (b), all /t/ tokens (c), and all /d/ tokens (d) as produced by Speaker 3 in monosyllables.

Before turning to the separate distributions of the disyllable tokens, Figure 6c above warrants more discussion. The unflapped /t/-tokens fall into a bimodal distribution in Figure 6c, because unflapped /t/ measured shorter in the second context, ranging from 60 – 180 ms in the carrier sentence (Figure 7b), than in context one (Figure 7a), ranging from 150 – 290 ms in the pseudorandomized list, and in context three (Figure 7c), ranging from 180 – 360 ms in the minimal pairs list. Figure 7 below illustrates that although unflapped /t/ tokens are shorter in Context 2 than Contexts 1 and 3, a visible gap still occurs after the 50 ms label in Figure 7b. Because words in sentences are shorter than words in isolation, it is not surprising that unflapped /t/ was shorter in this context.

Figure 7. Distribution of all /t/-tokens in Context 1 (a), all /t/-tokens in Context 2 (b), and all /t/-tokens in Context 3 (c) as produced by Speaker 3 in both monosyllables and disyllables.

Regardless of this shift in the distribution, a clear gap between the flapped and unflapped /t/-tokens in monosyllables still occurs after 52 ms across context in Figure 6c and by context as seen in Figures 7a - 7c. As can be seen in Figure 8a, a larger gap

Figure 8. Distribution of all /t/ and /d/ tokens (a), all /t/ tokens (b), and all /d/ tokens (c) as produced by Speaker 3 in disyllables.

occurs between the flapped and unflapped disyllable tokens, where the only unflapped tokens occur above 80 ms. Figures 8b and 8c illustrate that the unflapped disyllable tokens were all underlyingly /t/-tokens, and not /d/-tokens.

The benefits of defining flaps in this manner include that the duration at which /t/ and /d/ are considered flapped is based on the speaker's productions, taking speaking rate and any other speaker variation into account. This process also reflects that alveolar flaps as a class are shorter in duration than alveolar stops, but it does not imply any other relationship between flaps and stops; whereas, using a value like half the duration of an alveolar stop or one standard deviation below the mean of an alveolar stop assumes a connection between the durations of flaps and stops. The flap cut-off value used in this study is superior to other values used in previous studies because it is a nonarbitrary value based on a natural gap in each speaker's distribution of /t/ and /d/ tokens, allowing flaps to be defined on a speaker-by-speaker basis.

3.3 Disyllables

Flap frequency

Using the speaker-by-speaker values listed in Table 5 above, each intervocalic /t/ or /d/ utterance was defined as flapped or unflapped. Of the 6720 disyllabic tokens recorded, 49 or 0.7% were excluded due to pronunciation or recording errors. Of the other 6671 flappable words and nonwords, 88% were pronounced as flaps. As predicted, word medial-/d/ was flapped more frequently than word medial-/t/, and the 23% difference shown in Figure 9 below proved significant in a paired samples t-test [$t(3315) = 30.92, p < .0001$].

Figure 9. Flap frequency of /t/ and /d/ in disyllables.

A 2×3 (Gender $\times$ Context) repeated measures Analysis of Variance was conducted on flap output and main effects of Gender and Context were found. Females flapped more frequently with a mean flap output of 89% than males with 86% [$F(1, 3310) = 13.392, p < .0001$]. As seen in Figure 10 below, the difference was due to the percentage of /t/-tokens flapped since 99% of the /d/-tokens were flapped for both genders, contributing to an interaction between flap output and Gender [$F(1, 3310) = 13.035, p < .0001$]. It is important to point out that the interaction between Gender and Context was nonsignificant, so females flapped more frequently than males across all contexts [$F(1, 3310) = .837, p = .433$]. Note also that these results contradict the findings of Sharf (1960), Zue and Laferriere (1979), and Byrd (1994) who concluded that male speakers produced flaps more frequently than females. The different results could be due to the smaller number of participants in the Sharf (1960) and Zue and Laferriere (1979) studies, two (one male and one

Figure 10. Flap frequency of /t/ and /d/ by gender.

female) and six (three males and three females) participants, respectively. In addition, the Byrd (1994) study included nasal flaps, where ‘minty’ must be pronounced [ˈmɪɹi] in order to be considered a flap, but the present study is based on only intervocalic oral flaps.

The main effect of Context [$F(2, 3310) = 5.804, p = .003$] is illustrated in Figure 11 below, where it can be seen that /t/ is flapped most frequently in the pseudorandomized list context, the first context read and the context in which the participants were the least aware of the /t/ and /d/ contrast in flappable minimal pairs like ‘writer’ and ‘rider’. The second most frequently flapped context was the carrier sentence, which was read second, and the least frequent flapping occurred in the

minimal pairs context, the last context read and the one in which the speakers would have been most aware of the /t/ and /d/ contrast. While the flap frequency of /t/ is higher in carrier sentences than in minimal pairs, a Bonferroni post hoc analysis determined that /t/ was flapped significantly more often in the list context than both

Figure 11. Flap frequency of /t/ and /d/ by context.

the carrier sentence ($p = .043$) and the minimal pairs ($p = .003$) context, but that there were no significant differences in flapping between the carrier sentence list and the minimal pairs list. Although all three contexts did not significantly differ from one another in flap frequency, it is clear that speakers flapped /t/ more frequently when they were less aware of the /t/ and /d/ contrast due to the pseudorandomized nature of

the list than the second carrier sentence context, when they had already seen the words in a list and may have become aware of the contrast, and the third paired context, when they focused on the contrast due to the opposition of minimal pairs. There was also an interaction between flap output and Context, because the flap frequency of the /t/-disyllables varied according to context, but the flap frequency of /d/-disyllables remained near 100% regardless of context [$F(2, 3310) = 5.932, p = .003$].

Next, using only the tokens read in the minimal pairs context, a 2×2 (Gender $\times$ Order) repeated measures Analysis of Variance was carried out on flap frequency to determine whether the order of the list, if the list were read /t/-token first or /d/-token first, had affected flap frequency. Similar to the results reported above, there was an interaction between flap output and Gender, where females flapped the /t/-tokens more often than the males, 76% and 71% respectively, but the /d/-tokens were flapped 100% regardless of gender [$F(1, 1110) = 4.363, p = .037$], but the main effect of Gender was only nearing significance [$F(1, 1110) = 3.730, p = .054$]. This reflects the generalization discussed above: females flapped /t/-tokens more frequently than males.

In addition to the above interaction between flap output and Gender, there were also interactions between Gender and Order [$F(1, 1110) = 64.099, p < .0001$] and Gender, Order, and flap output [$F(1, 1110) = 63.291, p < .0001$]. As discussed above, the interaction with flap output occurred because /t/-words and nonwords varied in flap frequency, but /d/-words and nonwords did not. As shown in Figure 12

below, the female speakers flapped /t/-disyllables more frequently when the minimal pairs were read /d/-first, but the male speakers exhibited the opposite pattern, flapping /t/-disyllables more frequently when the minimal pairs were read /t/-first.

According to the Patterson and Connine (2001) study, high frequency and monomorphemic words are more likely to be flapped than low frequency and

Figure 12. Flap frequency of /t/-disyllables read in minimal pairs by order.

bimorphemic words. In order to test whether word frequency affected flapping, the Kučera-Francis (1982) written word frequencies for the disyllable lexical items were used to divide the words into two groups, high frequency and low frequency, based on a mean cut-off of 25. After determining that the mean word frequency of all disyllable lexical items was 25, lexical items with word frequencies lower than 25 were labeled low frequency words and lexical items with word frequencies at and

above 25 were labeled high frequency. An independent samples t-test was conducted comparing the flap output of the high frequency words to that of the low frequency words. The high frequency disyllables were flapped 88%, and the low frequency disyllables were flapped 89%, a nonsignificant difference [$t(5021) = .306, p = .759$]. In another independent samples t-test, the words were divided into two different groups, very low frequency words (0 words per million) and very high frequency words (100 words per million and above), to see if an effect of word frequency could be obtained. The very low frequency disyllables were flapped 84%, and the very high frequency disyllables were flapped 85%, another nonsignificant difference [$t(1491) = .321, p = .748$]. To see if the number of morphemes in a lexical item affected flap frequency, an independent samples t-test was used to compare the flap output of monomorphemic lexical items to bimorphemic ones. Words comprised of one morpheme were produced as flaps 88% while words with two morphemes were flapped 89%, a nonsignificant difference [$t(5021) = .294, p = .769$]. Based on these results, it seems fair to conclude that neither word frequency nor morphological complexity of lexical items affected how often they were flapped.

Finally, using the 28 disyllabic words that do not exist in minimal pairs with their 28 nonword counterparts, as shown previously in Table 2, a paired samples t-test was performed comparing the flap frequency of lexical items to that of nonwords. As shown in Figure 13 below, the lexical items were flapped 5% more often than the nonwords, a significant difference [$t(1644) = 4.323, p < .0001$]. Additionally, in a 2×2 (Gender $\times$ Lexical Status) repeated measures Analysis of Variance, in which

Lexical Status represents whether the disyllables are lexical items or nonwords, the interaction between Gender and Lexical Status proved nonsignificant [$F(1, 1639) = .001, p = .944$]. This indicates that the difference in flap frequency between lexical items and nonwords occurred across all speakers regardless of gender and that it was not driven by the higher flapping frequency of female speakers.

Figure 13. Flap frequency of lexical items and nonwords.

The above flap frequency results suggest that females, contrary to past research, flap disyllabic /t/-tokens more frequently than males, although it is unclear why males and females produced the minimal pairs differently depending on the order in which the pairs were presented. Flap frequency was also affected by context, so the more aware speakers are of a contrast between flappable /t/ and /d/, the less likely they are to flap /t/. And although word frequency and morphological complexity of lexical items do not affect how often they are flapped, lexical items are flapped more frequently than nonwords.

Measures of flap duration

Using only flapped disyllabic tokens, a paired samples t-test was carried out comparing the mean duration of /t/-flaps to that of /d/-flaps. As reported in prior studies, /d/-flaps measured longer in duration (30.4 ms) than /t/-flaps (29.5 ms), a difference in duration in the opposite direction of /t/ and /d/ unflapped stops, where the voiceless alveolar stops were longer than the voiced. However, unlike previous studies, the small difference between /t/ and /d/-flaps was significant [$t(2511) = 4.363, p < .0001$], probably due to the large sample size.

A 3×2 (Context $\times$ Gender) repeated measures Analysis of Variance was conducted on flap duration, and main effects of flap duration, Context and Gender were found. As Figure 14 below illustrates, the flaps in the carrier sentences were

Figure 14. Flap duration of /t/ and /d/ by context.

significantly longer in duration than those in the pseudorandomized list [$F(2, 2500) = 3.847, p = .021$]. The significance of this difference was verified by a Bonferroni post

hoc test [$p = .008$]. At first blush, this seems illogical, since words read in a sentence should be shorter overall than words read in a list, including the flaps within those words. However, looking at Table 6 below, it is clear that the longer mean flap duration in the carrier sentence context was only produced by the female speakers. Although the average duration of flaps was 0.4 ms longer in the carrier sentence context than in the list context for males, the females produced a larger difference (1.5 ms) than the male participants. Furthermore, there is an interaction between context and gender [$F(2, 2500) = 8.528, p < .0001$], because the female speakers produced the flaps in the carrier sentences with a longer duration than the males.

Table 6. Mean flap duration by gender and context.

Gender	Context	Mean (ms)	Gender	Context	Mean (ms)
female	List	31.7	male	list	27.2
	Carrier	33.2		carrier	27.6
	Pair	33.4		pair	26.0

In addition to a main effect of Gender [$F(1, 2500) = 421.926, p < .0001$], there was also an interaction between Gender and flap duration [$F(1, 2500) = 20.320, p < .0001$]. The main effect was due to females producing longer flaps (32.8 ms) than males (26.9 ms) as mentioned above, and the interaction, shown in Figure 15 below, occurred because females produced longer flaps for /d/-disyllables than /t/-disyllables. In fact, the main effect of flap duration, where /d/-flaps had a significantly longer duration than /t/-flaps, appears to be produced only by the female speakers.

Figure 15. Flap duration of /t/ and /d/ by gender.

Some of the flap duration findings, such as that the /d/-flaps were significantly longer than the /t/-flaps and that the mean flap duration in the carrier sentence context was longer than in the pseudorandomized list context, appear to contradict previous research, but those differences are only produced by a subset of speakers. The longer flap durations of the female speakers were probably due to their slower speaking rate; an independent samples t-test comparing the word length of flapped disyllables produced by female speakers to that of male speakers showed that female speakers spoke significantly more slowly (540.1 ms) than the male speakers (476.6 ms) [$t(5854) = 20.488, p < .0001$].

Using only the minimal pairs context, a 2×2 (Order $\times$ Gender) repeated measures Analysis of Variance was performed on flap duration to see if the order in

which the flappable minimal pairs were read, /t/-first or /d/-first, affected the mean duration of the flaps. Due to the slower speaking rate of the female speakers as explained above, a main effect of Gender was again found. The female speakers produced flaps with a mean duration of 33.3 ms, which was significantly longer than the mean duration of flaps produced by the male speakers (25.8 ms) [$F(1, 807) = 262.553, p < .0001$].

Although the main effect of Order did not reach significance [$F(1, 807) = 3.303, p = .070$], there was an interaction between flap duration and Order. As can be seen in Figure 16 below, when the minimal pairs were read /t/-first, the mean /t/-flap duration was shorter, but when the list was read /d/-first, the mean /d/-flap duration was shorter [$F(1, 807) = 76.031, p < .0001$]. This seems to be a list effect, in which the second item read in the list was longer in duration than the first, causing the second flap produced to also be longer.

Figure 16. Flap duration of /t/ and /d/ by order.

Although the flap duration of /d/-flaps appears to be longer than /t/-flaps, there are no significant differences in flap duration that are consistently exhibited across gender. Like previous studies, it must be concluded that flap duration is not a cue to the underlying voicing of /t/ or /d/-flaps. While overall flap duration differences between /t/ and /d/ fall below a 1 ms difference, speaker variation does exist as shown in Table 7 below. For example, speakers 14 (7 ms) and 18 (6 ms) exhibit a larger mean difference than the other speakers. Future perceptual studies could shed light on whether or not this larger difference is perceived by listeners as a cue for voicing.

Table 7. Speaker Variation in Flapped consonant duration of disyllables.

Speaker #	/t/ Duration (ms)	s.d.	Minimum	Maximum	/d/ Duration (ms)	s.d.	Minimum	Maximum	Difference (ms) (/t/ - /d/)
1	31	9.7	9	55	32	9.7	12	56	-1
2	33	9.3	12	57	36	9.4	12	56	-3
3	27	7.9	12	43	26	7.6	12	46	1
4	40	6.0	27	60	41	7.4	15	58	-1
5	33	10.5	10	59	33	10.7	10	57	0
6	27	7.5	14	56	26	7.2	12	48	1
7	26	6.3	17	49	26	7.1	14	42	0
8	24	5.3	11	41	24	5.8	14	41	0
9	29	9.8	14	54	29	8.2	15	53	0
10	25	5.2	12	40	25	5.5	14	42	0
11	35	9.0	10	57	36	9.5	11	57	-1
12	33	8.7	15	57	36	8.8	15	59	-3
13	28	8.8	10	46	31	8.6	12	51	-3
14	25	8.8	10	56	32	10.4	11	59	-7
15	42	16.1	16	66	39	10.9	18	59	3
16	27	7.8	13	45	27	7.1	14	45	0
17	27	7.5	14	51	28	8.0	15	55	-1
18	28	7.2	16	44	34	7.6	18	46	-6
19	27	7.8	18	54	28	7.2	18	52	-1
20	32	8.9	15	56	30	7.2	16	54	2

Measures of vowel duration

Of the five measurements recorded for each disyllabic token, four of them reflect vowel duration: first vowel duration, first vowel duration to second vowel duration ratio, word length, and first vowel duration to word length ratio. All four measures will be reported during the paired samples t-tests, because these measurements replicate the methods used in past studies; however, only the first vowel duration will be reported in the repeated measures of Analysis of Variance as the results of the other three related variables produced nearly identical results.

Paired samples t-tests were conducted on the four measurements mentioned above to see if first vowel duration varied depending on whether the disyllabic word contained a /t/-flap or a /d/-flap. As predicted, the vowel duration measures were significantly longer in all cases when the disyllabic token contained a /d/-flap. Figure 17 below shows that vowels preceding /d/-flaps were 8 ms longer than those

Figure 17. First vowel duration preceding /t/ and /d/-flaps.

preceding /t/-flaps, a significant difference [$t(2511)=-6.649$, $p<.0001$]. Table 8 below details the significance of the other vowel measurements. The results replicated past findings that first vowel to second vowel ratios, word lengths, and first vowel to word ratios are longer for tokens with /d/-flaps than with /t/-flaps.

Table 8. Significant vowel length measures in /t/ and /d/-flap

Measurement	/d/-flap disyllables	/t/-flap disyllables	difference /d/ - /t/	level of significance
V1:V2	1.75	1.67	.08	[$t(2497)=-4.867$, $p<.0001$]
Word Length	511 ms	501 ms	10 ms	[$t(2511)=-6.465$, $p<.0001$]
V1:V2	.279	.269	.01	[$t(2511)=-4.061$, $p<.0001$]

A $3 \times 2 \times 2$ (Context $\times$ Gender $\times$ Diphthong) repeated measures Analysis of Variance was conducted on vowel duration. In addition to the main effect of vowel duration explained above, there were also main effects of Context, Gender and Diphthong but no interactions. The main effect of Context, shown below in Figure 18, is an expected result [$F(2, 2500) = 44.545$, $p < .0001$]. Since the first context was a pseudorandomized list, the words and nonwords from that context would be read more slowly than items in the other two contexts, resulting in longer vowel duration. Similarly, the words and nonwords read in the third context, the minimal pairs list, would be read faster than the isolated tokens from the first context, but more slowly than the tokens read in the carrier sentence in the second context. Bonferroni post hoc tests revealed that all three contexts, the pseudorandomized list, the carrier

Figure 18. First vowel duration by context.

sentences, and the minimal pairs, differed significantly from one another in vowel length at the $p < .0001$ level.

Like the main effect of Context, the main effects of Gender and Diphthong should come as no surprise. Since the different speaking rates of the male and female speakers has already been well established, it is predictable that the female speakers produced significantly longer first vowels (137 ms) than the male speakers (131 ms) [$F(1, 2500) = 16.285, p < .0001$]. Just as predictably, disyllables containing diphthongs in the first syllable were produced with vowels that were 40 ms longer on average than those in words or nonwords containing monosyllables as illustrated in Figure 19 [$F(1, 2500) = 561.921, p < .0001$].

Figure 19. First vowel duration for monophthongs and diphthongs.

Contrary to predictions made based on previous research in perception, the interaction between vowel duration and Diphthong proved to be nonsignificant [$F(1, 2500) = .010, p = .998$]. As illustrated in Table 9 below, the vowel difference between monophthongs preceding /t/ and those preceding /d/ did not differ significantly from

Table 9. Nonsignificant interaction of vowel duration and occurrence of monophthongs and diphthongs.

Vowel Type	Preceding /d/ (ms)	Preceding /t/ (ms)	Difference (/d/ - /t/) (ms)
Monophthong	118	111	7
Diphthong	157	150	7

the vowel difference between diphthongs preceding /t/ and those preceding /d/. So vowel durational differences preceding /t/ and /d/-flaps were not significantly more striking for diphthongs (Sharf 1960) or monophthongs (Malécot and Lloyd 1968).

And, though the context in which the disyllables were read did affect the speaking rate of the participants, it had no effect on the vowel durational differences.

A further analysis was carried out in which monomorphemic disyllabic pairs, like ‘petal’ and ‘pedal’, were compared to bimorphemic disyllabic pairs, like ‘wetting’ and ‘wedding’. The purpose of this analysis was to see if vowel duration differences preceding flapped /d/ and /t/ would be greater for bimorphemic pairs than for monomorphemic pairs since the bimorphemic pairs are derived from root words, in this case ‘wet’ and ‘wed’, which contrast in vowel duration preceding /t/ and /d/. A paired samples t-test found that the difference in vowel duration for bimorphemic pairs, an 8 ms difference where vowels were longer preceding flapped /d/, was greater than that of monomorphemic pairs, a 3 ms difference where vowels were longer preceding flapped /d/, and that this difference was nearing significance [$F(1, 364) = 3.247, p = .072$]. These results are based on a small sample of stimuli, three monomorphemic pairs and five bimorphemic pairs, so a larger set of stimuli matched in morphemic structure could lead to clearly significant results.

Similar to speaker variation in flapped consonant duration discussed above, speakers also varied in vowel duration as shown in Table 10 below. Whereas some speakers, like speaker 5 (-2 ms) and 17 (1 ms), produced a smaller difference than the 8 ms average, other speakers, like speakers 13 (11 ms), 15 (10 ms), and 18 (28 ms), produced a larger than average difference. These speaker variations could be used in future perceptual studies to ascertain whether listeners use these consonant and vowel duration differences to decide if a flapped alveolar stop is a /t/ or /d/ or

Table 10. Speaker variation in vowel duration for disyllables

Speaker #	Vowel Duration preceding /r/ (ms)	s.d.	Minimum	Maximum	Vowel Duration preceding /ɛ/ (ms)	s.d.	Minimum	Maximum	Difference (ms) (/ɛ/ - /r/)
1	137	37.4	30	211	144	39.8	44	232	7
2	131	39.8	47	212	136	42.9	35	216	5
3	117	33.8	45	198	123	37.7	49	223	6
4	152	42.8	56	248	156	44.3	53	269	4
5	133	41.7	50	278	131	42.7	47	265	-2
6	122	32.6	51	198	129	35.4	58	220	7
7	129	37.2	48	196	134	39.3	62	212	5
8	132	35.7	38	215	138	36.7	43	222	6
9	118	31.6	67	182	127	39.1	61	205	9
10	110	31.0	33	170	115	33.0	32	195	5
11	130	40.0	48	213	133	41.0	43	224	5
12	126	36.7	43	204	133	40.8	47	232	8
13	118	38.5	45	205	129	41.2	29	214	11
14	149	41.0	64	241	158	41.2	63	245	9
15	110	33.7	73	176	120	44.0	70	193	10
16	140	36.0	61	250	147	36.3	68	230	7
17	110	31.7	24	184	111	33.7	32	209	1
18	140	36.1	68	210	168	50.9	67	291	28
19	145	36.7	38	242	148	36.5	61	261	3
20	130	33.4	34	232	138	34.8	55	233	8

whether they rely more heavily on word frequency or a general /d/-bias to make these judgments.

4. General Discussion

Instead of using arbitrary numerical cut-offs, like 10 – 40 ms or 50 ms, to distinguish the consonant duration of stops and flaps, the present cut-off value was determined on a speaker-by-speaker basis by visually analyzing the flap and stop distribution of each speaker. Using the flap frequency of speakers based on these values, it was found that females flapped /t/-tokens more frequently than males, in direct contradiction with the findings of Sharf (1960), Zue and Laferriere (1979), and Byrd (1994). As discussed briefly above, the findings of this study are more reliable because they are based on a larger number of speakers, 10 males and 10 females, a larger number of flappable tokens, 112 disyllabic tokens read in three contexts, and a more well-defined flapping environment, intervocalic post-stressed, than the previous research.

The context in which the flappable tokens were read also affected the frequency that the tokens were pronounced as flaps. When speakers read words and nonwords in the first context, a pseudorandomized list within which the words were read in isolation and the speakers were unaware of the contrast being investigated, participants were more likely to flap intervocalic /t/ and /d/. In the second context, a pseudorandomized carrier sentence list during which the speakers may have become aware of the contrast since they were seeing the list for the second time, the participants flapped to a great degree, though the difference was not significant.

Finally, in the third context, a minimal pairs list within which the words were set in contrastive /t/ and /d/ opposition, making the participants hyperaware of the contrast, flapping occurred the least and significantly differed from the list context. Therefore, the more aware the speakers were of the /t/ and /d/ contrast, the less likely they were to flap.

Based on the results reported above, it is clear that the participants in this study contrast voicing in monosyllabic words ending in /t/ and /d/ by using vowel duration, with the vowels preceding the voiced alveolar stop measuring significantly longer, and word-final consonant duration, with the voiceless alveolar stops measuring longer in duration. However, the use of vowel duration and consonant duration to distinguish disyllabic tokens is less clear. While it is true that the /d/-flaps measured significantly longer than the /t/-flaps, a 0.9 ms difference could not be used by a listener as cue for voicing. Furthermore, as mentioned above, this difference was produced only by the female participants who spoke more slowly than the male speakers. Since the difference was not produced by both male and female speakers, it is unlikely that flap duration is used to contrast voicing. Similarly, a vowel duration difference preceding /t/-flaps and /d/-flaps, where vowels measured 7.3 ms longer preceding /d/-flaps, was proven statistically significant. While this difference appeared consistently, pointing to an incomplete neutralization of /t/ and /d/ when flapped, it is also unlikely that listeners could use a 7.3 ms difference in preceding vowel length to distinguish /t/-flaps from /d/-flaps. The comparison of vowel duration differences between monomorphemic and bimorphemic disyllables also

produced interesting results, indicating that any vowel duration contrasts that exist preceding flapped /t/ and /d/ could be due to a relationship between monosyllable root words ('wet' and 'wed') and bimorphemic disyllabic flapped words ('wetting' and 'wedding'). Since a contrast is maintained through vowel duration in the corresponding root word, the contrast is likewise maintained in the corresponding bimorphemic word, paralleling a phonological output-output correspondence in Optimality Theory.

In order to evaluate whether /t/ and /d/ have been perceptually neutralized when flapped, a perceptual study will be necessary. If the /t/ and /d/ flaps are contrastive due to vowel and flap duration, we would expect a correlation between longer vowel duration and the perception of /d/-flaps and between longer flap duration and the perception of /t/-flaps. If /t/ and /d/ are neutralized perceptually, higher word frequency and a general /d/-bias, not acoustic measures, will be correlated with the correct perception of /t/ and /d/.

References

- Boermsa, Paul and David Weenink. 2005. Praat: doing phonetics by computer (Version 4.3.12) [Computer Program]. Retrieved May 11, 2005, from <http://www.praat.org/>.
- Byrd, Dani. 1994. Relations of sex and dialect to reduction. *Speech Communication* 15. 39 – 54.
- Charles-Luce, Jan. 1997. Cognitive factors involved in preserving a phonemic contrast. *Language and Speech* 40. 229 – 248.
- Charles-Luce, Jan and Kelly M. Dressler. 1999. The effects of semantic predictability in non-pathological older adults' production of a phonemic contrast. *Clinical Linguistics and Phonetics* 13. 199 – 217.
- Charles-Luce, Jan, Kelly M. Dressler and Elvira Ragonese. 1999. Effects of semantic predictability on children's preservation of a phonemic voice contrast. *Journal of Child Language* 26. 505 – 530.
- Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22. 129 – 159.
- Fisher, William M. and Ira J. Hirsh. 1976. Intervocalic flapping in English. *Chicago Linguistics Society* 12. 183 – 198.
- Fox, Robert A. and Dale Terbeek. 1977. Dental flaps, vowel duration and rule ordering in American English. *Journal of Phonetics* 5. 27-34.
- Francis, W. Nelson and Henry Kučera. 1982. Frequency analysis of English usage. Boston, Houghton Mifflin Company.

- Lehman, W. P. and R-M. S. Heffner. 1940. Notes on the length of vowels (III).
American Speech 15. 377 – 380.
- Malécot, André and Paul Lloyd. 1968. The /t/ : /d/ distinction in American alveolar
flaps. Lingua 19. 264 – 272.
- Patterson, David and Cynthia M. Connine. 2001. Variant frequency in flap
production: a corpus analysis of variant frequency in American English flap
production. Phonetica 58. 254 – 275.
- Sharf, Donald J. 1960. Distinctiveness of ‘voiced T’ words. American Speech 35.
105 – 109.
- Sharf, Donald J. 1962. Duration of post-stress intervocalic stops and preceding
vowels. Language and Speech 5. 26 – 30.
- Zue, Victor W. and Martha Laferriere. 1979. Acoustic study of medial /t,d/ in
American English. Journal of the Acoustical Society of America 66. 1039 –
1050.

Appendix A – Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens by speaker.

Full distribution of /t/ and /d/ tokens

Zoomed in to flap cut-off value portion
of the distribution

Speaker 1 (57 ms flap cut-off value marked by vertical black line)

Speaker 2 (61 ms flap cut-off value marked by vertical black line)

Appendix A Continued – Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens by speaker.

Full distribution of /t/ and /d/ tokens

Zoomed in to flap cut-off value portion
of the distribution

Speaker 3 (52 ms flap cut-off value marked by vertical black line)

Speaker 4 (61 ms flap cut-off value marked by vertical black line)

Appendix A Continued – Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens by speaker.

Full distribution of /t/ and /d/ tokens

Zoomed in to flap cut-off value portion
of the distribution

Speaker 5 (61 ms flap cut-off value marked by vertical black line)

Speaker 6 (58 ms flap cut-off value marked by vertical black line)

Appendix A Continued – Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens by speaker.

Full distribution of /t/ and /d/ tokens

Zoomed in to flap cut-off value portion
of the distribution

Speaker 7 (53 ms flap cut-off value marked by vertical black line)

Speaker 8 (43 ms flap cut-off value marked by vertical black line)

Appendix A Continued – Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens by speaker.

Full distribution of /t/ and /d/ tokens

Zoomed in to flap cut-off value portion
of the distribution

Speaker 9 (57 ms flap cut-off value marked by vertical black line)

Speaker 10 (43 ms flap cut-off value marked by vertical black line)

Appendix A Continued – Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens by speaker.

Full distribution of /t/ and /d/ tokens

Zoomed in to flap cut-off value portion
of the distribution

Speaker 11 (60 ms flap cut-off value marked by vertical black line)

Speaker 12 (60 ms flap cut-off value marked by vertical black line)

Appendix A Continued – Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens by speaker.

Full distribution of /t/ and /d/ tokens

Zoomed in to flap cut-off value portion
of the distribution

Speaker 13 (53 ms flap cut-off value marked by vertical black line)

Speaker 14 (60 ms flap cut-off value marked by vertical black line)

Appendix A Continued – Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens by speaker.

Full distribution of /t/ and /d/ tokens

Zoomed in to flap cut-off value portion
of the distribution

Speaker 15 (69 ms flap cut-off value marked by vertical black line)

Speaker 16 (47 ms flap cut-off value marked by vertical black line)

Appendix A Continued – Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens by speaker.

Full distribution of /t/ and /d/ tokens

Zoomed in to flap cut-off value portion
of the distribution

Speaker 17 (57 ms flap cut-off value marked by vertical black line)

Speaker 18 (59 ms flap cut-off value marked by vertical black line)

Appendix A Continued – Distribution of all monosyllabic and disyllabic /t/ and /d/ tokens by speaker.

Full distribution of /t/ and /d/ tokens

Zoomed in to flap cut-off value portion
of the distribution

Speaker 19 (55 ms flap cut-off value marked by vertical black line)

Speaker 20 (59 ms flap cut-off value marked by vertical black line)