American English Flapping: Evidence Against Paradigm Uniformity With Phonetic Features

Anastasia K. Riehl
Cornell University, Ithaca, NY, USA
E-mail: akr6@cornell.edu

ABSTRACT

This study investigates the claim that flapping patterns in American English are subject to phonetic paradigm uniformity constraints based on the phonetic feature [extra short closure], as proposed by Steriade [1]. The results of this study reveal that speakers do not maintain uniform paradigms with regard to flapping and that [extra short closure] is only one of several acoustic cues relevant for differentiating a flap from its corresponding stop.

1. INTRODUCTION

This study investigates the claim that flapping patterns in American English (AE) are subject to phonetic paradigm uniformity constraints based on the phonetic feature [extra short closure], as proposed by Steriade [1]. This claim asserts that the /t/ allophone a speaker articulates in the base form of a word (whether a flap or a stop) will be maintained in the inflected form, even if the phonology predicts a different outcome. I investigate this claim by subjecting the preliminary study in [1] to a larger, more controlled perceptual and acoustic study. The results of this investigation reveal that, a) flap/stop alternations cannot be explained by appealing to paradigm uniformity and that b) [extra short closure] is only one of several acoustic cues relevant for distinguishing a flap from a stop. These results, therefore, do not lend support for a collapse of the phonetic and phonological components of the grammar, as argued in [1].

1.1 Background

AE flapping has typically been regarded as a categorical phonological rule whereby /t/ and /d/ become a flap before an unstressed vowel (see e.g. [2]). The rule is considered optional, applying in fast and casual speech. More recently, several phonetic studies have attempted to predict the presumed optionality of flapping by investigating whether or not the occurrence of a flap can be attributed to prosodic factors. For example, [3] provides some evidence that prosodic context, such as the degree of stress and accent on the surrounding vowels, can affect the articulation of a /t/ such that a flap surfaces. For discussion of the various phonological and phonetic accounts of flapping, see [4].

A further hypothesis is that flapping patterns are a result of paradigm uniformity constraints [1]. The concept of paradigm uniformity, sometimes referred to as paradigm regularity or analogy, has long been a part of the phonological literature (see e.g. [5], also [6] and references therein for recent work). A uniform paradigm refers to a group of words that share a morpheme which prefers to be invariant across forms despite differences in phonological context. [1] extends paradigm uniformity to the concept of "phonetic analogy." The claim is that uniform paradigms exist at the phonetic level as well, such that an allomorph will surface in unexpected positions simply to satisfy a paradigm uniformity condition. On the basis of two case studies, [1] further concludes that since non-contrastive phonetic attributes can be subject to paradigm uniformity conditions, these attributes should not be regarded as distinct from phonological features. This argument is ultimately used as evidence for a collapse of the phonetic and phonological components of grammar.

In the case of AE flapping, [1] claims that the allophone of /t/ a speaker articulates in the base form of a word is the allophone that will surface in the inflected form, even if the phonology predicts a different outcome; thus, someone who pronounces *primitive* with a [t^h] due to secondary stress on the final syllable will pronounce the first /t/ in *primitivistic* as a [t^h], even though the environment for flapping exists in the latter form. [1] claims the difference between a flap and [t^h] is due to a non-categorical feature, [extra short closure], following phonetic studies in [7] and [8] that find flap durations to be shorter than stop durations.

1.2 Preliminary study in [1]

For the study in [1], 12 subjects read two lists of words. The first list contained five target words: v*oluntary*, *positive, negative, primitive* and *relative,* where the expectation was that some speakers would place secondary stress on the syllable following the /t/ and thereby produce a stop and that other speakers would not place secondary stress on this syllable and thereby produce a flap. These words were randomized with five others where all speakers were expected to flap: *fatal, fetish, notary, rotary,* and *totem,* (to ensure that stops were not being articulated artificially). On a second list, all 10 words were inflected with the productive suffix *–istic.* As there is presumably no secondary stress on the syllable following /t/ in the inflected forms, the phonology predicts that the speakers would pronounce these forms with a flap. However, [1] predicts that a speaker who pronounces the base form with a stop will also pronounce the inflected form with a stop, due to paradigm uniformity effects, even in the absence of secondary stress on the following syllable. The results of the study in [1] appear to support the claims of paradigm leveling, although there is a small amount of variation in the data that is not accounted for. Eleven out of 12 speakers

 ISBN 1-876346-48-5

have identical allophones for each pair of base and inflected form. One speaker contains one non-uniform pair.

The study in [1] produced interesting results but was subject to methodological limitations, as speakers repeated each word only once, and tokens were not subject to acoustic analysis. The rationale for the present study was to test the important hypothesis in [1] through a more quantitative investigation.

2. METHODOLOGY

2.1 Recordings

Six subjects were recorded for this experiment, three males (Sp1, Sp2 and Sp3) and three females (Sp4, Sp5 and Sp6), all between the ages of 26 and 32, and all native speakers of AE from the northeastern United States. The subjects were asked to read 24 lists of words, based on the list in [1] as follows: half of the lists contained the five target base forms, the five flap forms, and five forms in which subjects were expected to use a stop allophone (*atomic, attentive, protective, retentive, fantastic*). The stop forms did not appear on the wordlist in [1] but were added to maintain a balance between flapped and stopped forms, so that the speakers did not get into a pattern of flapping the target words simply due to influence from the other flapped forms on the list. The other half of the lists contained the inflected forms of the target words and the inflected forms of the flapped words (all with the suffix *–istic* attached) as well as the stopped forms (uninflected). The lists were randomized, and filler words were added to the beginning and end of each, in an effort to offset effects of list intonation. The two sets of lists were then intermingled such that a list of base forms was followed by a list of inflected forms, and so on.

Before the recordings, the subjects, who were all naive as to the purpose of the experiment, were given an opportunity to familiarize themselves with the wordlists. Digital recordings were then made in the soundproof booth in the Cornell Phonetics Laboratory. Each recording session took place in six short blocks, totaling approximately 25 minutes for each speaker. A subject was given four lists at a time. After reading the lists, the subject was given a chance to take a break if desired. Most breaks were approximately two minutes long, with none longer than five minutes.

2.2 Perception test

Three listeners completed a perceptual test to identify occurrences of stops and flaps in the data. The listeners were all graduate students at Cornell University with a background in phonetics and native speakers of AE, one of these subjects being the author. The target tokens were divided into six sets, each set containing the repetitions of a single speaker. The sequence of tokens in each set was randomized, and a Perl script was created to run on each set, whereby the program would play a sound file twice and then, after prompting from the listener via pressing a key on the keyboard, the next token would be repeated twice, and so forth. The listeners heard the tokens through headphones, and on a sheet of paper checked a box indicating whether they heard a stop or flap, in a forced choice experimental paradigm. After each set, the listeners were asked to take a break (which ranged from 10 minutes to 20 hours), with all sets completed within a 24 hour period. Before beginning each set, the listeners were given a tutorial which included a practice test of four to six tokens from the relevant speaker.

2.3 Acoustic analysis

The recordings were sampled at a rate of 22050 Hertz and then labeled and analyzed using ESPS Xwaves. The labels were assigned as follows. The beginning of the closure was placed at the point where both the second formant (F2) and the third formant (F3) of the preceding vowel ended, or, in the cases where the formants remained throughout the closure, the label was placed at the point where the intensity of F2 and F3 decreased. The end of the closure was placed at the onset of a burst, or in cases where no burst was present, at the point where F2 and F3 of the following vowel began or increased in intensity.

Fundamental Frequency (F0) was then calculated using the ESPS utility *getf0*, associated with Xwaves, at a step size of one millisecond (ms), and a Perl script was used to place a VOT label at the first point (of at least five consecutive points) where the probability of voicing was one, starting at the closure label. The script calculated the duration of the closure, VOT and the total duration (closure + VOT). An additional script calculated the percentage of voicing in the closure and in the second half of the closure.

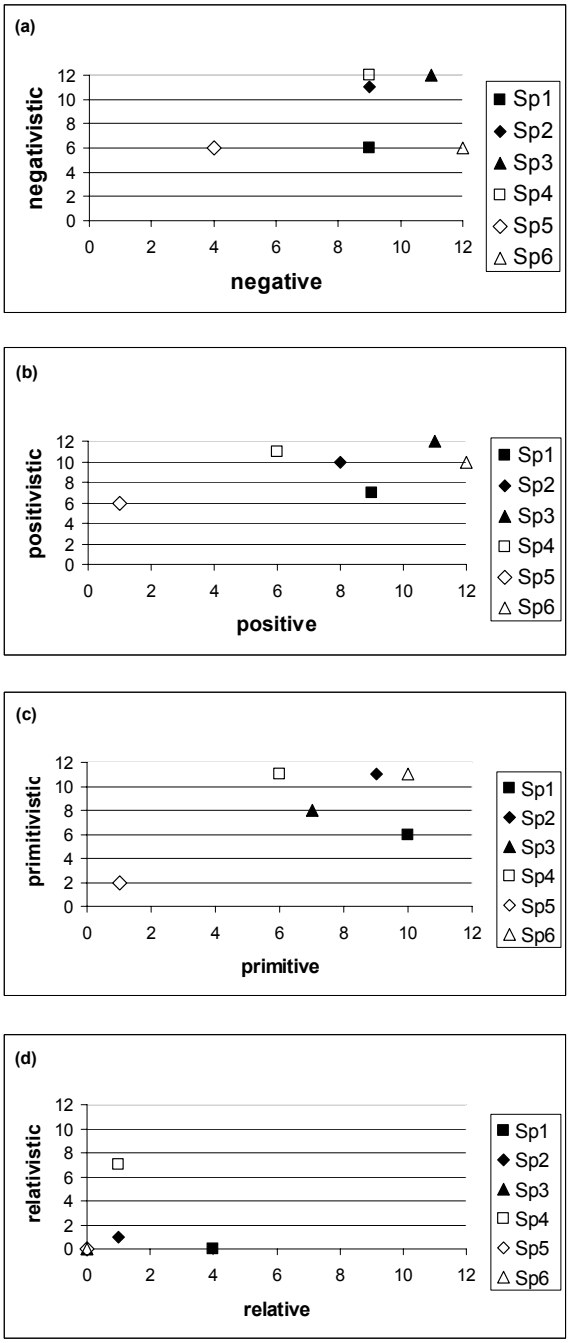
All 12 repetitions of each of the target base forms and each of the inflected forms were analyzed for each speaker. The pair *voluntary/-istic* is not included in this discussion, since all six of the speakers in this study were unable to articulate a flap in the environment following a /n/. (Also note that all 12 of the subjects in [1] pronounced these words with a stop as well.) With four target base forms and four target inflected forms repeated 12 times each by each of six subjects, the total number of tokens to be discussed in this study is 576.

3. RESULTS

3.1 Perception test

Of the 576 target tokens played for the three listeners, all three agreed on the categorization of the /t/ allophone in 543 of the cases, leaving disagreement in only 5.7% of the cases. The following figures report the listeners' categorization of the allophones in each of the four target words. In cases of tokens that do not have a unanimous classification, the majority judgment is reported (i.e. a token is considered a stop if two of the three listeners identified it as a stop). Since a uniform paradigm refers to a speaker maintaining a consistent allophone across the repetitions of a base and inflected form, such a paradigm can be identified in the figures as a value at point 0-0 or 12-12, meaning that a speaker uttered all stops or all flaps, respectively, all 12 times. A value anywhere else on a figure represents a non-uniform paradigm.

ISBN 1-876346-48-5



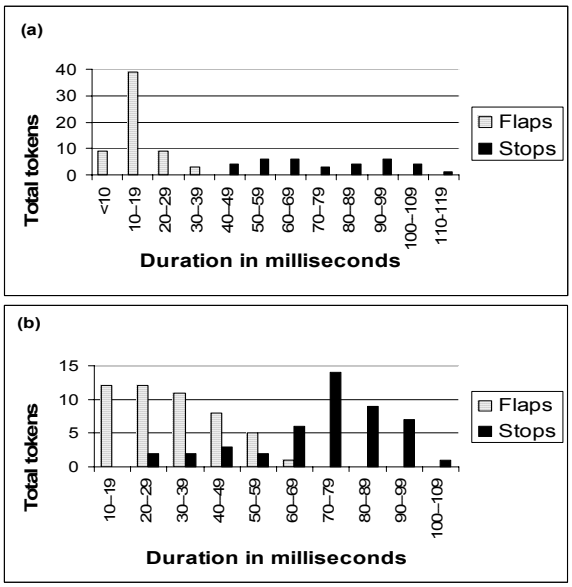
Figures 1a-d: Number of perceived flaps in base forms (x-axis) plotted against number in inflected forms (y-axis) across all speakers for (a) *negative/-istic*, (b) *positive/-istic* (c) *primitive/-istic*, and (d) *relative/-istic*.

As seen in Figures 1a-c, the *negative/-istic*, *positive/-istic* and *primitive/-istic* pairs exhibit similar patterns. No speaker maintains a perfectly uniform paradigm in any of these pairs. Several speakers do come close, however: in (1a) *negative/-istic*, Sp3 articulates 11 base and 12 inflected flaps; in (1b) *positive/-istic*, Sp3 articulates 11 base and 12 inflected flaps, and Sp6 articulates 12 base and 10 inflected flaps; and in (1c) *primitive/-istic*, Sp6 articulates 10 base and 11 inflected flaps. In the remaining cases (14 of the 18 pairs), the speakers articulate a mixture of flaps and stops across the members of a paradigm (between base and inflected) as well as within repetitions of a single form.

The pattern in (1d) *relative/-istic* is quite different. Three speakers (Sp3, Sp5, Sp6) articulate perfectly uniform paradigms with no occurrences of a flap in either the base or inflected forms. Additionally, Sp2 comes close with one flap surfacing in each of the base and inflected forms. The other two speakers (Sp1 and Sp4) articulate a mix of flaps and stops. Interestingly, all speakers articulate a majority of stops in *relative/-istic*, whereas only Sp5 maintains a majority of stops in the other three target word pairs.

3.2 Duration

Total durations of the /t/ allophones in all of the target words were calculated as described in Section 2.3. The duration values of allophones perceived as flaps were compared to those perceived as stops for each speaker. The following Figures 2a-b display the results for Sp 2 (2a) and Sp1 (2b). (Note the different scales). Results for the other speakers are not included due to space limitations but are summarized below and reported in [4].



Figures 2a-b: Total durations of /t/ allophones perceived as flaps compared to those perceived as stops for Sp2 (a) and Sp1 (b).

As seen in Figure 2a, Sp2's flap and stop tokens form two distinct groups, with flaps ranging from 0 to 39 ms and peaking between 10 and 19 ms, and stops falling in a fairly even distribution between 40 and 119 ms. Sp1's tokens, however, seen in Figure 2b, do not form two discrete groups. Flap durations fall within the 10 to 69 ms range, with a majority of tokens between 10 and 29 ms, while stop durations fall within the 20 to 109 ms range, with a majority between 70 and 79 ms. There is some degree of overlap

ISBN 1-876346-48-5

between the flap and stop tokens, from the 20 to 69 ms range. Results for the other four speakers can be summarized as follows. Sp4 shows a degree of overlap between flap and stop tokens similar to Sp1. Sp3, Sp5 and Sp6 all exhibit a clearer division between flaps and stops, but with a small degree of overlap at the border.

3.3 VOT and closure voicing

VOT and percent of voicing during the closure were also calculated for each speaker's flap and stop allophones, as described in Section 2.3. The voicing results discussed are from the second half of the closure rather than the full closure, as these better capture the voicing of the consonant by avoiding voicing from the preceding vowel. Though lack of space prohibits a detailed presentation of these results, they are described in [4] and summarized in Figure 3, where A indicates no overlap between flaps and stops with regard to the relevant cue, B indicates an overlap of one to five tokens, and C indicates an overlap of more than five tokens.

	Sp1	Sp2	Sp3	Sp4	Sp5	Sp6
Duration	C	A	B	C	B	B
VOT	C	A	B	C	B	A
Voicing	C	B	B	C	C	B

Figure 3: Overlap between flaps and stops with regard to total duration, VOT, and percent of voicing in the second half of closure, for each speaker. A=no overlap, B=overlap of 1 to 5 tokens, C= overlap of more than five tokens.

As seen in Figure 3, duration and VOT values both relate perfectly to flap/stop identification for Sp2, and VOT correlates perfectly for Sp6. For Sp5, duration and voicing both correlate well, though not perfectly, and for Sp1, Sp3 and Sp4, all three cues correlate equally well, with Sp1 and Sp4 having the poorest matches in all three categories.

4. DISCUSSION

4.1 Paradigm uniformity

As seen above in Section 3.1, of the 24 base-inflected pairs examined (6 speakers x 4 word pairs), only three exhibit perfectly uniform paradigms, and only five more come close (where at least 10 of 12 repetitions of each of the base and inflected form are consistent). The majority, 16 pairs, exhibit variation rather than uniformity in the articulation of the allophone, both across forms and within repetitions of the same form. Further, the three cases of perfectly uniform paradigms and one of the cases of a near-uniform paradigm all involve the pair *relative/-istic*, implying that some characteristic of this particular pair, rather than a generalization about uniform paradigms, is responsible for the consistency in articulation of the stop allophones.

These results are contrary to those presented in [1], where 59 out of 60 word pairs (12 speakers x 5 words) were reported to exhibit uniform paradigms. However, in that study, each speaker was asked to read each word only once, not allowing for variation across multiple repetitions, which was found to play a significant role in the present study. The question of what factors may be responsible for the variation observed in this study is addressed in [4]. However, the data presented here are sufficient to conclude that paradigm uniformity is not responsible for the observed patterns.

4.2 Duration, VOT and voicing

The results of this study reveal that there is a relationship between duration and an allophone's flap/stop identification. The relationship is perfect for Sp2 and near-perfect for Sp3, Sp5 and Sp6. However, there is some overlap for most speakers, and considerable overlap for two speakers (Sp1 and Sp4). An investigation of VOT and closure voicing reveals that these cues also relate to an allophone's flap/stop identification. For some speakers, one or more of these cues correlate better than or as well as duration. The assumption in [1], that [extra short closure] is the relevant cue for distinguishing flaps from stops, is therefore probably an oversimplification: duration does appear to be a significant cue in the identification of a /t/ allophone, but it is not the only, and perhaps not the most important, one. Ongoing work is exploring how cue weighting (as proposed in [9]) may play a role in the perception of flaps.

5. CONCLUSION

The results of this study offer evidence against a paradigm uniformity account of flapping patterns in AE and against [extra short closure] as the sole cue for flap identification. This study therefore does not provide evidence for the claim in [1] that AE flapping patterns support a collapse of the phonetic and phonological components of grammar.

ACKNOWLEDGEMENTS

I am grateful to Abigail Cohn, Amanda Miller-Ockhuizen, and Draga Zec for feedback on previous drafts of this paper; Janet Pierrehumbert and Ayako Tsuchida for helpful discussions; and Eric Evans for technical support.

REFERENCES

[1] D. Steriade, "Paradigm uniformity and the phonology-phonetics boundary." Paper presented at the 5th Conference in Laboratory Phonology, Evanston, Illinois, 1996.

[2] D. Kahn, *Syllable-based Generalization in English Phonology*. New York: Garland, 1980.

[3] K. DeJong, "Stress-relative variation in the articulation of coda alveolar stops: flapping revisited." *Journal of Phonetics* 26, pp. 283-310, 1998.

[4] A. Riehl, "American English flapping: Perceptual and acoustic evidence against paradigm uniformity with phonetic features." To appear in *Working Papers of the Cornell Phonetics Lab* 15, 2003.

[5] P. Kiparsky, "Historical linguistics," in *A Survey of Linguistic Science*, W. Dingwall, Ed., pp. 577-649. College Park, MD: Linguistics Program, University of Maryland, 1971.

[6] J. McArthy, "Optimal paradigms." ROA-485, 2002.

[7] V. Zue & M. Laferriere, "Acoustic study of medial /t, d/ in American English." *JASA* 66, pp. 1039-1050, 1979.

[8] S. Banner-Inouye, *Trills, Taps and Stops in Contrast and Variation*. UCLA dissertation, 1996.

[9] G.A. Miller & P.E. Nicely, "Analysis of perceptual confusions among some English consonants." *JASA* 27, p.p. 338-53, 1955.