The American English Flapping Rule
and the Effect of Stress on Stop Consonant Durations

Alice Turk

1. Introduction

The experiment described in this paper concerns the American English Flapping Rule, whereby non-word-initial intervocalic /t/ and /d/ are flapped when preceding an unstressed vowel, as in the words 'métal' and 'pyramidal'; in contrast the /t/ preceding a stressed vowel in 'mégalic' is not flapped. Acoustically, a major difference between flapped and non-flapped alveolars is durational: flaps are considerably shorter than their non-flapped counterparts. In this study we measured the durations of intervocalic stops in both the non-flapping and flapping environments as well as the durations of vowels that preceded the alveolars in order to answer the following questions:

1. Does a general timing mechanism underly the rule? It may be the case that the extreme shortening found in /t/ and /d/ is an instantiation of a more general timing principle. We discuss possible mechanisms for the length differences between poststress and prestress stops as well as implications for the formulation of the American English flapping rule.

2. Are vowels preceding flaps that were underlying /d/’s longer than vowels preceding flaps that were underlying /t/’s? It is well-known that vowels preceding voiced stops tend to be longer than vowels preceding voiceless stops (e.g. Chen, 1970). Longer vowels before flapped /d/’s than before flapped /t/’s would suggest that lengthening before voiced segments is due to the underlying phonological distinction ([+voice] vs. [-voice]) rather than to the phonetic realization of the consonant (which, in the case of most flaps, is as a voiced segment). In fact, one study (Fox and Terbeek, 1977) found that vowels preceding flapped /d/’s were longer than vowels preceding flapped /t/’s.

3. Does a longer phonological phrase affect segment durations?

2. The Phonology of Flaps

Flaps are allophones of /t/ and /d/ formed by a rapid movement of the tongue-tip making contact with the alveolar ridge, followed by immediate release. They have characteristically short durations of 10-40 ms. (Zue and Laferriere, 1979). In American English, /t/ and /d/ are realized as flaps in the following environments:

103
1. After a stressed vowel and before an unstressed vowel:
\[ v \_ v \] For example: mé/t/āl \rightarrow me[D]al ( [D] is used as the symbol for a flap).

2. Optionally between two unstressed vowels:
\[ v \_ v \] For example: provoca/t/īve \rightarrow provoca[D]īve or provoca[t]īve

3. Optionally word-finally between two vowels (stressed or unstressed):
\[ v \_ #v \] For example: no/t/ā/t/̣ all \rightarrow no[D] a[D] all or no[t] a[t] all.

priva/t/e airplane \rightarrow priva[D] airplane

Word-initial alveolars are never flapped.

The experiment described in this paper is less concerned with the phonological features involved in the flapping rule than with a phonetic analysis of the phenomenon. However, the phonological accounts of the flapping phenomenon are relevant because they all refer to prosodic structure in the formulation of the rule, suggesting that all consonants occurring in the flapping environments should exhibit the same phonetic consequences of being in that particular environment. These accounts vary in the segmental environment they claim is relevant, the distinctive features with which they characterize the flap, the domain of the application of the rule, and in the prosodic environment of flapping. From a prosodic point of view, the accounts can be divided into three principle views: 1. Flaps are ambisyllabic (Kahn, 1976; Gussenhoven, 1986); 2. They are syllable-final (Selkirk, 1982; Inouye, 1989); 3. They are non-foot-initial (Kiparsky, 1979). These views are described below.

2.1 Flaps are ambisyllabic

Kahn (1976) and Gussenhoven (1986) both claim that the occurrence of alveolar flaps can be predicted by referring to the syllabic structure notion of ambisyllabicity. It is described as follows for word-medial consonants (Kahn, 1976:55,56):

Ambisyllabicity

\[
\begin{array}{c}
\text{In \ [-cons] C C0 V} \\
\text{S1 \quad S2 \quad -stress}
\end{array}
\]

Associate C and S1

Kahn describes four allophones of the voiceless alveolar stop:

1. Voiceless aspirated alveolar stop: [th] ex: creativity
2. Unreleased voiceless alveolar/glottal stop: [tʰ] ex: hat
3. Released voiceless alveolar stop: [t] ex: stem
4. Voiced alveolar flap: [D] ex: item

Their distribution can be predicted from the syllable structure described above:
Stops which are both syllable-initial and non-syllable-final are aspirated.
Alveolar stops which are syllable-final and non-syllable-initial are glottalized.
Stops which are non-syllable-initial and non-syllable-final are released and not aspirated.
Alveolars which are both syllable-initial and syllable-final (ambisyllabic alveolars) are flapped.

The rule for flapping (ordered after the rule for aspiration) can be written as follows:

/θ/ -> [D] / [-cons] a<#(>)> +syllabic
                b <-stress>

Conditions:
1. ~a -> b (if the alveolar stop is word-medial, the following vowel must be unstressed for flapping to occur).
2. if a, within connected phrases only.
3. Does not apply in artificially slow speech.

Kahn (1976) assigns [-cons] to prevent the alveolar stops in "after", "faster", and "kept it" from flapping.

Kahn hesitates as to which feature to assign to the flap itself: in order to capture the voicing characteristic of many flaps, he assigns [-stiff v.c]; and to be consistent with feature descriptions of trilled /r/’s which are considered taps, he assigns the flap the feature specification of [+sonorant].

2.2 Flaps are Syllable-final

Selkirk (1982) differs from Kahn (1976) principally in that in her account, the consonants Kahn describes as ambisyllabic are resyllabified in order to be the coda of the preceding syllable. She cites the following evidence for resyllabifying the consonants including those which are flapped: the lengthening of vowels before voiced consonants, the nasalization of preceding vowels, and l-velarization, based on the premise that these are syllable-conditioned processes. Selkirk claims that resyllabified consonants behave as though they were underlyingly syllable final: it is the property of syllable finality that accounts for the flapping phenomenon.

Selkirk's syllabification rules have different implications for the prediction of alveolar stop allophones in American English. Instead of having to specify that aspiration applies to stops which are syllable initial and not syllable final, her rule is simplified in that all syllable-initial stops are aspirated. To distinguish syllable-final alveolars which are to be realized as flaps from those which are to be realized as glottalized alveolars, she posits the
feature Release as the determining parameter. Alveolars which are [-release] are glottalized; alveolars which are [+release] are flapped:

From Selkirk (1982):

Tap (= Flap)

t,d -> D / s(...[-cons] ___ )s

| +release

Glottalization

p,t,k -> p', t', k' / ___

| -release

Selkirk claims that [+release] is the unmarked value for the feature and describes at length the environments for [-release], although, as Inouye (1989) points out, she does not unify the environments for [-release] in any principled way. Although Selkirk states that the feature Release is not known to be phonemically contrastive in any language (and that perhaps it is not a distinctive feature, she notes that different languages have different realizations of Release for their stop consonants (e.g. French stops are always released except when followed by a homorganic consonant; in Vietnamese, prepausal consonants are not released), and that Release must be specified at some level of representation.

Another account which treats flaps as syllable-final is that of Inouye (1989). Her paper deals mainly with the issue of the phonological specification of the flap, however, which has been virtually ignored in the accounts of Kahn (1976) and Selkirk (1982). She argues that the flap is phonologically dynamic, and proposes that the flap is actually a contour segment (cf. Sagey (1986)), based on evidence that the flap exhibits edge effects (the lefthand environment ≠ the righthand environment). On the left edge, flapping is reported to be blocked by [+cons] segments (that have a constriction in the oral cavity); on the right edge, flapping seems to be blocked by segments which require the tongue tip to be above a certain threshold. Her argument rests on the commonly held observation that flapping does not occur in words such as "after", "faster", and in phrases such as "kept it" (the left edge is [+cons]). It also does not occur when the right edge is [+coronal], as in "beaten". She argues from such evidence that the dynamic nature of the flapping gesture has not been captured either in the feature specification of Kahn (1976) ([+son], [-stiff v.c]), or in Selkirk's requirement that only alveolar stops which are [+release] be flapped. Inouye proposes a multivalued feature Aperture with which to characterize the approach,
contact and release stages of the flap, and describes the flapping rule as one of spreading of
the aperture node from adjacent segments onto the coronal segment.

2.3 Flaps are Non-foot-initial

Kiparsky (1979) differs from the rest in that his analysis does not refer to syllable
structure in describing the flapping environment. Kiparsky claims that the rules for
alveolars are foot-sensitive: only non-foot-initial alveolars which follow a [-cons] segment
and precede a [+syllabic segment] may flap. For example:


Basically, within a foot, alveolar stops that precede unstressed vowels can be flapped:
re[pe[D]i[D]ive]S

To summarize, scholars are divided as to what is the prosodic environment of
flapping. Kahn and Gussenhoven claim that flaps are ambisyllabic; Selkirk and Inouye
claim flaps are syllable-final; and Kiparsky argues that they are non-foot-initial. The results
of this experiment will be discussed in terms of which prosodic account they support.

3. The Phonetics of Flapping

The study draws on the observation of Zue and Laferriere (1979) that
characteristically flaps are very short (10-40 ms); the experiment described here seeks to
determine whether this characteristic short duration is a product of the stress on
surrounding segments: i.e., whether it is a phonetic consequence of being in that prosodic
environment.

Several studies have dealt with the effect of stress on the durations of intervocalic
consonants.

Zue and Laferriere (1979) in a study of /t/ and /d/ in American English found that
word-medial alveolars are realized as extremely short flaps when immediately following a
stressed vowel or when between two unstressed vowels. They report that flaps between
two unstressed vowels are significantly longer than flaps immediately following a stressed
vowel.

As for other places of articulation, Umeda (1977) and Stathopoulos and Weismer
(1983) found that closure durations of both voiced and voiceless stops immediately
preceding a stressed vowel were longer than that of stops immediately following a stressed
vowel. Umeda's study consisted of one speaker reading a twenty minute essay. Umeda
notes that velars show the smallest range of difference, although no tests of statistical
significance are reported. /b/ was not included in the prestress environment, nor were total stop durations reported for the voiceless stops. Stathopoulos and Weismer (1983) report similar results listed in the following table:

<table>
<thead>
<tr>
<th></th>
<th>Mean Durations of stop closures in 2 positions:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.prestress</td>
</tr>
<tr>
<td>b</td>
<td>92 s.d. 10</td>
</tr>
<tr>
<td>p</td>
<td>96 s.d. 12</td>
</tr>
<tr>
<td>d</td>
<td>76 s.d. 13</td>
</tr>
<tr>
<td>t</td>
<td>82 s.d. 13</td>
</tr>
<tr>
<td>g</td>
<td>68 s.d. 10</td>
</tr>
<tr>
<td>k</td>
<td>72 s.d. 11</td>
</tr>
</tbody>
</table>

Closure durations for /b,p,d,t, and g/ seem to be significantly different in the two stress environments, (although they do not report statistical significance). Interestingly, for the non-alveolar segments, differences are larger in the voiced cases (/b,g/).

In another study, Davis and Summers (1989) report that in word-medial VC sequences, total durations (closure + aspiration) were longer in a prestress environment than in a poststress environment. However, the test tokens used for velars were not well-paired in terms of segmental environment--the post-stress velars were prevocalic, whereas the prestress velars occurred before /t/ (sagging and sacking vs. degrees and degrees). They did not show any statistical tests for the effect of stress on consonant duration.

In general, all of the above studies suggest that consonants preceding stress within a foot are longer than consonants that follow stressed vowels. However, a clearer picture of the effects of stress on both parts of intervocalic consonant durations (closure + aspiration) is needed. The experiment described in this paper was designed to provide a complete description of the effects of stress on word-medial stop consonants.

Of relevance to the second part of the experiment is the study by Fox and Terbeek (1977) on the durations of vowels preceding flaps. Fox and Terbeek conducted a study of 20 disyllabic word pairs (ex: writer, rider) spoken by 21 speakers. They discovered that stressed vowels preceding flapped /d/’s were significantly longer than stressed vowels preceding flapped /t/’s, suggesting that the lengthening before voiced stops rule makes reference to the phonological representation of segments, not to their phonetic realization.

However, this seems to be the case only for stressed vowels. Davis and Summers (1989) confirmed Fox and Terbeek’s results for stressed vowels, but they found that unstressed vowels preceding voiced prestress stops were longer, but not significantly so, than those preceding voiceless stops. In the experiment described below, the durations of both stressed and unstressed vowels preceding alveolar stops are reported.
4. The Experiment

The first part of this experiment was designed to determine systematically whether the stress effect on stop durations suggested by previous research can in fact be generalized to word-medial stop consonants of all places of articulation. Measurements of closure, aspiration and total (closure + aspiration) duration were taken from stops of all places of articulation in real words. The division of stops into closure and aspiration portions may show which part of the stop is shortened or lengthened. Two different sentence lengths were used to see if a longer phonological phrase had any effect on the duration of the test consonant. None was expected: Klatt (1976) observed that there is little evidence that speakers adjust segment durations in order to satisfy a global rhythmic constraint. However, an affect of sentence length on segment duration would indicate that segmental timing is dictated at least in part at a prosodic level higher than the word.

In the second part of the experiment, the duration of vowels preceding the alveolars in the different stress environments was measured to see if lengthening before stop voicing is observed for both types of flaps.

4.1 Methods
4.1.1 Stimuli

Closure, aspiration, and total durations of /p,t,k,b,d,g/ and the durations of vowels preceding alveolars were investigated when the test consonant occurred in three word-medial environments:

1) Before a stressed syllable (the non-flapping environment).
2) After a stressed syllable and before an unstressed syllable (a flapping environment).
3) Between unstressed syllables (Optional flaps: Zue and Laferriere found that flaps in this environment were longer than the other flaps)

The test words were chosen such that for a given word-medial consonant, the segments surrounding the stops were as similar as possible:

<table>
<thead>
<tr>
<th>v_V</th>
<th>V_v</th>
<th>v_v</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p/</td>
<td>regair</td>
<td>léper</td>
</tr>
<tr>
<td>/t/</td>
<td>megallic</td>
<td>métal</td>
</tr>
<tr>
<td>/k/</td>
<td>losâle</td>
<td>lócal</td>
</tr>
<tr>
<td>/b/</td>
<td>rebél</td>
<td>râbble</td>
</tr>
<tr>
<td>/d/</td>
<td>medâllion</td>
<td>médal</td>
</tr>
<tr>
<td>/g/</td>
<td>cigâr</td>
<td>vigor</td>
</tr>
</tbody>
</table>
The frame sentences (listed in Appendix A) were selected so that the test words would be in similar prosodic positions. Sentence length was increased by adding a modifier before each test word.

In all, there were 18 test words x 2 sentence lengths x 4 speakers x 5 repetitions. Alveolar data was collected for three out of the four speakers.

Test sentences were presented to the speakers along with an equal number of decoy sentences. The sentences were typed on notecards and were shuffled before each repetition of the set. They were read at a normal speaking rate, with natural intonation patterns in an IAC sound-treated booth.

4.1.2 Speakers

Two female and two male native speakers of American English served as speakers for the experiment. All were in their early twenties and speak a standard dialect of American English. Speaker AM (male) is a native of Ithaca, NY; Speaker AW (male) grew up in New York City; Speaker LE (female) is from California; Speaker RG (female) is a native of Ann Arbor, Michigan.

4.1.3 Measurements

The speech was digitized at 8 kHz, lowpass filtered at 4 kHz and stored on a SUN 3/160 workstation. Measurements were taken from waveforms and spectrograms displayed using AT&T Bell Labs WAVES software. The beginning of the stop was judged to be the point of F1 offset, and likewise, the end of the stop was the F1 onset. When F1 could not be seen clearly, measurements were taken at points of obvious change in the waveform where there was no higher formant energy in the spectrogram.

4.2 Results

A repeated measures ANOVA with factors of stress environment (Stress Position), underlying phonological voicing (Voicing), place of articulation (Place), sentence length (±Modifier) and the repeated measures factor of speaker (Speaker) shows main effects of Stress Position, Place, and Voicing on closure, aspiration and total consonant durations:

Effect of Stress on

- Closure: $F(2,144) = 20.660; p < .0001$
- Aspiration: $F(2,144) = 435.858; p < .0001$
- Total duration: $F(2,144) = 421.103; p < .0001$
Effect of Place of Articulation on
   Closure: F(2,144) = 315.866; p < .0001
   Aspiration: F(2,144) = 148.960; p < .0001
   Total duration: F(2,144) = 252.543; p < .0001

Effect of Voicing on
   Closure: F(1,144) = 18.47; p < .0001
   Aspiration: F(1,144) = 1858.35; p < .0001
   Total duration: F(1,144) = 1443.26; p < .0001

Effect of Speaker (Repeated Measures Factor) on
   Closure: F(3,432) = 227.03; p < .0001
   Aspiration: F(3,432) = 135.51; p < .0001
   Total duration: F(3,432) = 334.62; p < .0001

Sentence length had no significant effect on either the closure duration, the aspiration duration, or the total duration of the stops (p > .05). Interactions of the repeated measures factor of Speaker with all combinations of Stress Position, Voicing, and Place are significant at the p<.05 level, indicating that the different speakers behave differently. All interactions which include the factors of Stress Position, Voicing, and Place, but not ±Modifier are significant at the p<.05 level, which indicates that the effect of different stress environments varies according to the voicing and place specification of the stop.

Since the stress effects are confounded by the place of articulation and underlying voicing of each stop, repeated measures ANOVAs were conducted which tested the effect of Stress Position and ±Modifier on stop closure durations, aspiration durations and total durations for each type of stop. Likewise, single factor analyses of variance testing the effects of Stress Position on each type of measurement for each speaker were done to expose the nature of inter-speaker variability. Results broken down in this way are given below.

4.2.1 The Effect of Stress Position

4.2.1.a On Total Stop Durations

A repeated measures ANOVA including the factors of Stress Position and ±Modifier with the repeated measures factor of Speaker shows a main effect of Stress Position on the total durations (closure + aspiration) of all stops except /ɡ/. F and p values are listed in Appendix A. As can be seen from Figure 1, the difference in duration in different stress environments lies primarily between the non-flapping and the flapping environments. Repeated measures ANOVAs with factors of ±Modifier, Stress which
included only the two flapping environments, and the repeated measures factor of Speaker showed that total durations in the two flapping environments were significantly different only in the case of /d/ (F (1, 16) = 126.89; p < .0001) (Figure 1).

1. Pooled: Total Durations

Results from ANOVAs and Scheffe F-tests for each speaker follow. F and p values are given in Appendix C.

Speaker AW (Figure 2):

Bilabial stops and /k/ are significantly longer in the prestress (non-flapping) environment than in the two flapping environments. The effect of Stress on /g/ was in the predicted direction, although differences were not significant.

There is no significant difference between the closure durations in the two flapping environments.
2. Speaker AW: Total Durations

Speaker AM (Figure 3):

Total durations are significantly longer in the pre-stress environment only in the case of the alveolars. /k/ exhibits differences in the expected direction, although differences are not significant.

The durations of stops in the two non-flapping environments are significantly different only in case of /b,d, and g/. Zue and Lafriere (1979) had found that flaps which occur between two unstressed vowels are longer than those in the environment V_v. This is the case for /d/ in the present study, but /b/ and /g/ have the opposite pattern.

3. Speaker AM: Total Durations
**Speaker RG (Figure 4):**

Total durations are longer in the prestress environment for all stops except /g/. There is no significant difference between the two non-flapping environments for any stop, although the trend in all cases is for the stop between two unstressed vowels to be longer.

4. **Speaker RG: Total Durations**

![Graph showing total durations for /p/, /t/, /k/, /b/, /d/, /g/ in different environments: v, v-v, v-v-v.]

**Speaker LE (Figure 5):**

Total durations are longer in the prestress environment for all stops except /g/. Significant differences between the two flapping environments are found for /p,k,b and d/. /d and k/ are longer in the v-v flapping environment than in the V-v flapping environment.

5. **Speaker LE: Total Durations**

![Graph showing total durations for /p/, /t/, /k/, /b/, /d/, /g/ in different environments: v, v-v, v-v-v.]

To summarize, total durations of all stops except for /g/ are longer in the non-flapping environment than in the flapping environments for three out of four speakers.

The durational differences that Zue and Laferriere (1979) observed for alveolars in the two flapping environments (/t/ and /d/ tend to be longer in the environment v_v than in the environment V_v) were not observed for stops of other places of articulation, nor did the alveolars the present experiment consistently follow the pattern Zue and Laferriere observed.

In order to determine which component of each consonant (closure or aspiration, or both) is responsible for observed total duration differences, the effect of stress on both closure and aspiration are shown in the two sections which follow.

4.2.1.b On Closure Durations

In a repeated measures ANOVA including Stress Position, ±Modifier and the repeated measures factor of Speaker, Stress Position had a significant effect on the closure durations of /t,b, and d/. A significant effect in the opposite direction was found for /p/ -- its closure duration in the prestress environment is shorter than its closure duration in the flapping environments.

6. Pooled: Closure Durations

Speaker AW (Figure 7):

Closure durations for /b/ are significantly longer in the non-flapping environment than in both of the flapping environments. The closure durations for /k/ are significantly longer in the non-flapping environment than in the v_v environment, but the difference in the non-flapping environment and the V_v environment is not significant, although in the
expected direction. The differences shown for /p and g/ between the non-flapping and flapping environments are not significant.

As for durational differences between stops in the two flapping environments, /b/ is the only consonant whose closures are significantly longer in the v_v environment than in the V_v flapping environment.

7. Speaker AW: Closure

![Graph showing closure duration for different consonants in Speaker AW's speech]

Speaker AM (Figure 8):

Only the closure durations of the alveolars (/t and d/) are significantly longer in the prestress environment than in both of the flapping environments. The rest of the consonants (/p,k,b,g/) are longest in the V_v environment.

8. Speaker AM: Closure

![Graph showing closure duration for different consonants in Speaker AM's speech]
**Speaker RG** (Figure 9):

Closure durations for /t/, /b/, and /d/ are significantly longer in the prestress environment than in both flapping environments.

A significant difference in the expected direction between the two flapping environments was found only for /d/.

**9. Speaker RG: Closure**

![Closure Duration Graph](image)

**Speaker LE** (Figure 10):

Closure durations for /t/, /b/, and /d/ are longer in the prestress environment than in both of the flapping environments, although the difference between the closure duration of /d/ in the v\_V and v\_v environments was not significant. The closure durations of /p/, /k/ and /g/ are shorter in the prestress environment.
Only /d/ shows a significant difference in the expected direction between the two flapping environments.

10. Speaker LE: Closure

![Graph showing closure durations for different consonants](image)

To summarize, the closure durations of /t/ and /d/ are consistently longer in the prestress environment than in the flapping environments. As for consonants of other places of articulation, /b/ is the only one which is longer in the prestress environment for three out of four speakers.

4.2.1.c. On Aspiration Durations:

Stress Position had a significant effect on the aspiration durations of all stops except /g/ (Figure 11). The aspiration durations for /p/ were significantly different in the expected direction in the two flapping environments (F(1,16) = 23.97; p < .0002).
11. Pooled Aspiration

Speaker AW (Figure 12):

The aspiration durations are significantly longer in the prestress environment for /p, k,b, and g/. Only /p/ shows a significant difference between the two flapping environments.

12. Speaker AW: Aspiration
**Speaker AM** (Figure 13):

The aspiration durations are significantly longer in the prestress environment than in the other two environments for /p,t,k/. /p/ aspiration shows a significant difference between the two flapping environments.

13. Speaker AM: Aspiration

![Graph showing aspiration in ms for /p, t, k, b, d, g/](image)

**Speaker RG** (Figure 14):

The aspiration durations are significantly longer in the prestress environment for /p,t,k/.

Aspiration durations for /p and g/ are significantly longer in the v_v environment than in the V_v environment.
**14. Speaker RG: Aspiration**

![Graph showing aspiration durations for /p, t, k, b, d, g/ with different environments.]  

**Speaker LE (Figure 15):**

The aspiration durations are significantly longer in the prestress environment for /p, t, k, and d/. The aspiration durations for /k/ are significantly longer in the \_v \_v environment than in the V \_v environment.

**15. Speaker LE: Aspiration**

![Graph showing aspiration durations for /p, t, k, b, d, g/ with different environments.]  

To summarize, all speakers show greater aspiration durations in the prestress environment than in the other two environments for all voiceless stops.

The aspiration durations of /p/ are longer in the \_v \_v flapping environment than in the V \_v flapping environment for three out of four speakers.
4.2.2. The Effect of a Longer Phonological Phrase

For the pooled data, the main effect of ±Modifier is significant only in the cases of the aspiration and total durations of /d/ (F(1,24) = 12.80; p < .0015—aspiration; F(1,24) = 7.74; p < .0103) and in the aspiration durations of /k/ (F(1,24) = 6.205; p < .02). These results confirm the observation of Klatt (1976) that there is little evidence that speakers adjust segment durations in order to satisfy a global rhythmic constraint.

4.2.3. Durations of Vowels Preceding /t/'s and /d/'s

Data from three speakers were included in a repeated measures ANOVA with factors of Stress Position, Voicing, ±Modified, and the repeated measures factor of Speaker. It shows a significant main effect of consonantal Stress Position (F(2,48) = 585.97; p < .0001) on the duration of the vowel before /t/ or /d/, which confirms that stressed vowels are longer than schwa (Stressed vowels had a mean duration of 104 ms; unstressed vowels had a mean duration of 43 ms). There is a significant interaction of Speaker with Stress Position indicating that stress position affects the durations of vowels differently for each speaker.

There is no main effect of Voicing. However, there is a significant interaction of Stress Position and Voicing on the duration of vowels before /t/ and /d/ (F(2,48) = 19.42; p < .0001), which indicates that the effect of Voicing on the length of the preceding vowel varies with the position of stress (Figure 16).

16. Pooled: Vowel Durations before /t/, /d/

![Graph showing vowel duration]
To explore the interaction of Stress Position and Voicing further, repeated measures ANOVAs were conducted to determine the effect of Voicing on the durations of vowels preceding /t/ and /d/ for each consonantal stress environment. In the prestress environment, vowels preceding /t/ were significantly longer than vowels preceding /d/ (F(1,18) = 20.36; p < .0003), and there was a significant main effect of Speaker (F(2,36) = 7.09; p < .0025). In the V_v consonantal environment, vowels preceding /d/ were longer than vowels preceding /t/, but not significantly (F(1,18) = 1.38; p > .2561). There was a significant interaction with Speaker (F(2,36) = 25.23; p < .0001). In the v_v
c environment, vowels preceding /d/ were significantly longer than vowels preceding /t/ (F(1,18) = 18.67; p < .0004), and again, there was a significant interaction with Speaker (F(2,36) = 3.85; p < .0305).

Mean durations of the pooled data for vowels preceding /t/ and /d/ in each stress environment are given below:

<table>
<thead>
<tr>
<th></th>
<th>before /d/ ([+voice])</th>
<th>before /t/ ([−voice])</th>
</tr>
</thead>
<tbody>
<tr>
<td>vCV</td>
<td>n = 30</td>
<td>n = 30</td>
</tr>
<tr>
<td></td>
<td>45 ms.</td>
<td>56 ms.</td>
</tr>
<tr>
<td>VCv</td>
<td>n = 30</td>
<td>n = 30</td>
</tr>
<tr>
<td></td>
<td>106 ms.</td>
<td>101 ms.</td>
</tr>
<tr>
<td>vCV</td>
<td>n = 30</td>
<td>n = 30</td>
</tr>
<tr>
<td></td>
<td>43 ms.</td>
<td>29 ms.</td>
</tr>
</tbody>
</table>

Graphs of the results from each speaker follow; they mimic the patterns found in the pooled data.

17. Speaker AM: Vowel Duration before /t/, /d/

![Bar chart showing vowel duration before /t/ and /d/ for V_v, v_v, and v_V environments. The chart includes bars for [+voice] and [−voice] conditions.](image-url)
4.3. Discussion

Generalizations that can be made from the results are discussed below:

4.3.1 The Effect of Stress

All stops (except /g/) are longer in total duration in prestress position than they are in either of the flapping environments. This finding suggests that relative consonant length is an acoustic correlate of the prosodic structure required by formulations of the flapping rule: the extreme shortening found in the alveolar cases is an instantiation of a more general
timing phenomenon. The failure of /g/ to conform to the pattern could be a result of the relative inertia of the tongue dorsum as compared to other articulators.

The general pattern of the results bears upon the formulation of the prosodic domain of the American English flapping rule. There is considerable disagreement in the literature over what the prosodic environment for flapping is: Kahn (1976) and Gussenhoven (1986) claim that flaps are ambisyllabic; Selkirk (1982) and Inouye (1989) claim that flaps are syllable-final, and Kiparsky (1979) claims that flapping is not syllable-conditioned, but occurs when alveolars are non-foot-initial.

The present experiment shows that total duration is an acoustic correlate of the prosodic structure common to all stops in each stress environment. If we assume that flaps are ambisyllabic, following Kahn (1976), we must therefore say that ambisyllabic consonants are shorter than those that belong to only one syllable. Chierchia (1983/1986) argues that geminate consonants (presumably long) are ambisyllabic. The two accounts do not necessarily conflict, if it is the case that the mapping from prosodic phonology to the acoustics is language-specific. However, it seems counterintuitive to have such extremely different mappings in different languages.

In Selkirk's (1982) theory, word-medial consonants following stressed vowels are resyllabified and become the codas of the preceding syllable. Results of the present study show that in an algorithm mapping prosodic structure to acoustic values, syllable-final consonants would be assigned relatively short durations, while syllable-initial consonants would be assigned relatively long durations. The results indicate that there would have to be no difference in the durations assigned to syllable-final consonants in stressed vs. unstressed syllables. However, Statthopoulos and Weismer (1983) found that in fact, closure durations of word-final consonants in CVCVC nonsense disyllables were shorter when the final syllable was stressed than when it was unstressed. Their results show that the durations of syllable-final stops vary according to the stress assigned to the syllable. The fact that the present experiment shows no difference in stop duration in the two non-flapping environments (V_v) and (v_v), does not support Selkirk's theory, where the stops would be codas of syllables that differ in stress, and would be expected to have different durations.

In Kiparsky's (1979) formulation of the flapping rule, no resyllabification is required; non-foot-initial alveolars are flapped. The shortening of stops in the flapping environment could be explained in two ways: (1) The shorter stops are shorter because they are compensating for the length of the stressed vowel that occurs earlier in the foot. Such an explanation would be consistent with theories of isochrony, which seek to explain the
phenomenon of decreasing syllable duration with an increasing number of syllables in a word (Kozhevnikov and Chistovich (1965), Port (1980), Allen (1976) and Lehiste (1976). The unit of isochrony is unclear, but Lehiste (1970) feels that it is either the foot or the word. It is also unclear exactly how the compensation takes place. It is possible that shortening takes place within syllabic units; alternatively, compensation could occur on a segmental level without reference to syllable structure; (2) Assuming duration is assigned in syllabic units, the shorter stops could be shorter simply because they are onsets in unstressed syllables (V_v and v_v), while the longer stops are long because they are onsets in stressed syllables (v_V). The results of Stathopoulos and Weismer (1983) for word-initial stop closures support this view. In nonsense disyllables of the form CVCVC, closure durations of word-initial stops were shorter when the first syllable was unstressed than when it was stressed.

Of the three prosodic domains posited for the flapping rule, the foot of Kiparsky (1979) seems to have the most support from results of the experiment described here.

Although the total durations show a clear pattern of being longer in the prestress environment and shorter in the flapping environment, there is a dichotomy in closure duration differences for voiced and voiceless stops. Closure durations of voiced stops are significantly longer in prestress position than in poststress position, whereas the closure durations of voiceless stops are not significantly longer in prestress position. This is consistent with the results of Stathopoulos & Weismer (1983) who found that closure durations for non-alveolar voiceless stops did not exhibit much of a difference between prestress and poststress position durations, whereas closure durations for voiced stops were very different in prestress position vs. poststress position. However, Stathopoulos & Weismer (1983) found that, despite the dichotomy between voiced and voiceless stops, the closure durations for all stops were longer in prestress position, which is not the case here. This experiment shows that for the voiceless stops, the shortening/lengthening occurs primarily in the aspiration portion, whereas for the voiced stops, shortening/lengthening occurs primarily in the closure portion. Aspiration thus appears to be more compressible/stretchable than closure. Kozhevnikov and Chistovich (1965) and Klatt (1976) have noted that vowels are more stretchable/compressible than consonants. Our experiment supports their findings, and also suggests that aspiration could be added to the stretchability/compressibility hierarchy: Vowels are the most stretchable/compressible, followed by aspiration; closure durations are the least stretchable/compressible.
4.3.2. Vowels Preceding /t/ and /d/  
Vowels preceding flapped /d/’s are significantly longer than vowels preceding flapped /t/’s only when the vowel before the flap is unstressed. When the vowel is stressed, length differences are in the expected direction, but are not significant. When the alveolar is in the non-flapping environment, the unstressed vowels preceding /d/ are actually shorter than vowels preceding /t/. The results differ slightly from those of both Fox & Terbeek (1977) and Davis and Summers (1989). Fox and Terbeek (1977) found that stressed vowels before flapped /d/ were reliably longer than stressed vowels before /t/. They concluded that the durations of preceding vowels were conditioned by the phonological specification of voicing for the consonant. Our results, while consistent with their conclusion, do indicate that the magnitude of lengthening before [+voiced] flaps varies with the stress environment of the flap. Furthermore, the slightly different results suggest that dialectal/idioidialectal differences play a role in lengthening before voicing phenomena.

Davis and Summers (1989) found that vowels preceding prestress voiced stops tended to be longer (but not significantly) than vowels preceding prestress voiceless stops. The results in this experiment are in the opposite direction. Davis and Summers (1989) used their non-significant result to support the view of Maddieson (1985) that lengthening before voicing is not a syllable-conditioned process (the lengthened vowel is not tautosyllabic with the following consonant in the vCV case). The results of this experiment tend to support the view expressed in Selkirk (1982) that lengthening before voicing is a syllable-conditioned process. However, the most that can be tentatively concluded from the variety of results obtained here and across the literature is that there are dialectal/idioidialectal differences with respect to lengthening before voicing phenomena. Alternatively, it is possible that in some studies, observed effects are artifacts of the way the test sentences are presented.

4.3.3. The Effect of a Longer Phonological Phrase  
The presence of a modifier before the test word has virtually no effect on the durations in the test consonants for all speakers. These results corroborate Klatt’s (1976) observation that speakers tend not to adjust segment durations in order to satisfy global rhythmic constraints.
4.3.4. Interspeaker Variability

The stress effect described above as well as the lengthening before [+voiced] flaps varies according to speaker; speaker AM in this study showed significant length differences between the flapping and non-flapping environments only for the alveolars and /k/, while in other speakers the effect is more widespread.

5. Summary

The experiment has shown the following:

1) For all stops except /g/, a phonetic consequence of occurring in flapping environments is reduced length: Total stop durations word-medially in all places of articulation tend to be longer in prestress position. This suggests that the American English Flapping Rule is driven by an underlying timing mechanism. The results are slightly different than those of Stathopoulos and Weismer (1983) who found that stop closure durations tend to be longer in prestress position.

2) The results suggest that the American English Flapping Rule is a foot-conditioned process; they support the account of Kiparsky (1979) in which only non-foot-initial alveolar stops may flap.

3) Voiced stop shortening/lengthening takes place in the closure position, whereas voiceless stop lengthening/shortening takes place mostly in the aspiration portion. Comparing the closure, aspiration, and vowel duration differences in different stress environments shows a compressibility/stretchability hierarchy where vowels are most stretchable/compressible, aspiration is less compressible/stretchable, and closure is the least compressible/stretchable.

4) Unlike the results of Fox and Terbeek (1977), significant differences were found between measurements of vowels preceding flapped /t/’s and measurements of vowels preceding flapped /d/’s only when the alveolar occurred between two unstressed vowels.

Notes

1. In this study, closure durations of voiced stops are shorter than the closure durations of voiceless stops for all places of articulation in the poststress environment (V_v), and closure durations of voiced stops are shorter than the closure durations of voiceless stops for non-alveolars in the between unstressed environment (v_v). In contrast, in the prestress environment (v_V), voiced stop closures are longer than voiceless stop closures for labials and alveolars. In the velar case, voiced stop closures are shorter than voiceless stop closures. These findings support the findings of Davis and Surmmer...
that the direction of length differences between closures of voiced and voiceless stops in the prestress position is variable. Also, the findings support their claim that closure duration is a reliable cue to voicing after a stressed vowel.

References


Appendix A
He avoided the leper for health reasons.
He avoided the dirty leper for health reasons.
She preferred to repair the green car.
She preferred to thoroughly repair the green car.
He misplaced the caliper on his workbench.
He misplaced the outside caliper on his workbench.
He found the metal in the ground. (Speakers RG, AM and LE)
He found the copper metal in the ground. (Speakers RG, AM and LE)
Her dress looks metallic from a distance. (Speakers RG, AM and LE)
Her dress looks vaguely metallic from a distance. (Speakers RG, AM and LE)
She measured the diameter with a ruler. (Speakers RG, AM and LE)
She measured the indicated diameter with a ruler. (Speakers RG, AM and LE)
We expect the local on track fifteen.
We expect the incoming local on track fifteen.
She found the locale of the murder.
She found the mysterious locale of the murder.
He observed the follicle under the microscope.
He observed the little follicle under the microscope.
He watched the rabble from the tower.
He watched the cheering rabble from the tower.
She preferred to rebel against her captors.
She preferred to publicly rebel against her captors.
She studied the parable in religion class.
She studied the second parable in religion class.
She won the medal at the Olympics. (Speakers RG, AM and LE)
She won the silver medal at the Olympics. (Speakers RG, AM and LE)
He found the medallion on the floor. (Speakers RG, AM and LE)
He found the copper medallion on the floor. (Speakers RG, AM and LE)
The object looked pyramidal in shape. (Speakers RG, AM and LE)
The object looked perfectly pyramidal in shape. (Speakers RG, AM and LE)
He admired the vigor of the athletes.
He admired the steroid-induced vigor of the athletes.
He smoked the cigar in the waiting room.
He smoked the Cuban cigar in the waiting room.
He added the vinegar to the salad.
He added the cider vinegar to the salad.

Appendix B
1. The Effect of Stress on Total Durations
F and p values of the effect of Stress on total durations of /p, t, d, b, d, g/ from a repeated measures ANOVA with dependent variables of stress environment (Stress Position), sentence length (±Modifier), and the repeated measures factor of Speaker.

/p/: F (2,24) = 17.37; p < .0001
/t/: F (2,24) = 370.80; p < .0001
/d/: F (2,24) = 57.17; p < .0001
/b/: F (2,24) = 80.76; p < .0001
/d/: F (2,24) = 176.73; p < .0001
/g/: F (2,24) = .56; p < .5784

2. The Effect of Stress on Closure Durations
F and p values of the effect of Stress on closures durations of /p, t, d, b, d, g/ from a repeated measures ANOVA with dependent variables of stress environment (Stress Position), sentence length (±Modifier), and the repeated measures factor of Speaker.

/p/: F (2,24) = 54.27; p < .0001
/t/: F (2,24) = 17.14; p < .0001
/d/: F (2,24) = 2.423; p < .0001
/b/: F (2,24) = 61.49; p < .0001
/d/: F (2,24) = 167.31; p < .0001
/g/: F (2,24) = 1.47; p < .5784

3. The Effect of Stress on Aspiration Durations
F and p values of the effect of Stress on aspiration durations of /p, t, d, b, d, g/ from a repeated measures ANOVA with dependent variables of stress environment (Stress Position), sentence length (±Modifier), and the repeated measures factor of Speaker.

/p/: F (2,24) = 78.24; p < .0001
/t/: F (2,24) = 405.40; p < .0001
/d/: F (2,24) = 67.40; p < .0001
/b/: F (2,24) = 15.74; p < .0001
/d/: F (2,24) = 13.33; p < .0001
/g/: F (2,24) = 3.21; p < .5784

Appendix C
F and p values from one factor ANOVAs for each consonant for each speaker:

A. Stress on Total Consonant Duration
Speaker AW:
/p/: F(2,27) = 10.88; p < .0003
/t/: not collected

Speaker AM:
/p/: F(2,27) = 4.18; p < .0003
/t/: F(2,27) = 106.62; p < .0001
Results from a one factor ANOVA which compares the total durations of stops in the two flapping environments (v_v vs. v_v):

**Speaker AW:**

/p/: F(2,18) = 3.31; p > .0853

/t/: not collected

/k/: F(2,18) = 12.05; p < .0027

/b/: F(2,18) = 10.09; p < .0052

/d/: not collected

/g/: F(2,18) = .11; p > .7416

**Speaker AM:**

/p/: F(2,18) = 3.14; p > .0934

/t/: F(2,18) = 2.10; p < .1647

/k/: F(2,18) = 4.20; p > .0554

/b/: F(2,18) = 22.41; p < .0002

/d/: F(2,18) = 26.24; p < .0001

/g/: F(2,18) = 18.27; p < .0005

**Speaker RG:**

/p/: F(2,18) = .13; p > .7178

/t/: F(2,18) = .02; p > .8989

/k/: F(2,18) = 1.19; p > .2906

/b/: F(2,18) = .60; p > .4498

/d/: F(2,18) = 9.64; p < .0061

/g/: F(2,18) = 6.07; p < .024

**Speaker LE:**

/p/: F(2,18) = 7.31; p < .0145

/t/: F(2,18) = 1.39; p > .2545

/k/: F(2,18) = 27.46; p < .0001

/b/: F(2,18) = 9.85; p < .0057

/d/: F(2,18) = 50.33; p < .0001

/g/: F(2,18) = .32; p > .5769

**B. On Closure Durations:**

**Speaker AW:**

/p/: F(2,27) = .96; p > .3942

/t/: not collected

/k/: F(2,27) = 6.53; p < .0049

/b/: F(2,27) = 44.99; p < .0001

/d/: not collected

/g/: F(2,27) = .31; p > .7333

**Speaker AM:**

/p/: F(2,27) = 102.39; p < .0001

/t/: F(2,27) = 4.22; p < .0254

/k/: F(2,27) = 4.91; p < .0152

/b/: F(2,27) = 10.46; p < .0004

/d/: F(2,27) = 48.93; p < .0001

/g/: F(2,27) = 7.89; p < .002

**Speaker RG:**

/p/: F(2,27) = 15.54; p < .0001

/t/: F(2,27) = 11.41; p < .0003

/k/: F(2,27) = 5.89; p < .0075

/b/: F(2,27) = 18.32; p < .0001

/d/: F(2,27) = 28.55; p < .0001

/g/: F(2,27) = 2.43; p > .1068

**Speaker LE:**

/p/: F(2,27) = 16.03; p < .0001

/t/: F(2,27) = 5.04; p < .0138

/k/: F(2,27) = .21; p > .8136

/b/: F(2,27) = 16.59; p < .0001

/d/: F(2,27) = 65.14; p < .0001

/g/: F(2,27) = 4.24; p < .0251

Results from a one factor ANOVA which compares the closure durations of stops in the two flapping environments (V_v vs. v_v):
Speaker AW:
/p/: F(2,18) = 1.42; p > .2483
/t/: not collected
/k/: F(2,18) = 6.35; p < .0214
/b/: F(2,18) = 10.09; p < .0052
/d/: not collected
/g/: F(2,18) = .11; p > .7416

Speaker AM:
/p/: F(2,18) = 43.51; p < .0001
/t/: F(2,18) = 2.10; p > .1647
/k/: F(2,18) = 4.93; p < .0393
/b/: F(2,18) = 22.41; p < .0002
/d/: F(2,18) = 29.10; p < .0001
/g/: F(2,18) = 14.04; p < .0015

Speaker RG:
/p/: F(2,18) = 15.00; p < .0011
/t/: F(2,18) = 1.90; p > .1852
/k/: F(2,18) = .41; p > .5304
/b/: F(2,18) = .01; p > .9345
/d/: F(2,18) = 16.93; p < .0007
/g/: F(2,18) = 2.72; p > .1162

Speaker LE:
/p/: F(2,18) = 16.59; p < .0007
/t/: F(2,18) = .62; p > .4398
/k/: F(2,18) = .01; p > .9305
/b/: F(2,18) = 8.29; p < .01
/d/: F(2,18) = 65.72; p < .0001
/g/: F(2,18) = .23; p > .6354

C. On Aspiration Duration

Speaker AW:
/p/: F(2,27) = 23.30; p < .0001
/t/: not collected
/k/: F(2,27) = 8.09; p < .0018
/b/: F(2,27) = 14.58; p < .0001
/d/: not collected
/g/: F(2,27) = 8.32; p < .0015

Speaker AM:
/p/: F(2,27) = 80.89; p < .0001
/t/: F(2,27) = 192.68; p < .0001
/k/: F(2,27) = 19.44; p < .0001
/b/: F(2,27) = 1; p > .3811
/d/: F(2,27) = 2.29; p < .1207
/g/: F(2,27) = .99; p > .38

Speaker RG:
/p/: F(2,27) = 15.02; p < .0001
/t/: F(2,27) = 86.43; p < .0001
/k/: F(2,27) = 39.23; p < .0001
/b/: F(2,27) = 2.74; p > .0823
/d/: F(2,27) = 3.82; p < .0347
/g/: F(2,27) = 3.36; p < .0496

Speaker LE:
/p/: F(2,27) = 56.83; p < .0001
/t/: F(2,27) = 251.91; p < .0001
/k/: F(2,27) = 29.93; p < .0001
/b/: F(2,27) = 1.50; p > .24
/d/: F(2,27) = 13.28; p < .0001
/g/: F(2,27) = .60; p > .555

Results from a one factor ANOVA which compares the aspiration durations of stops in the two flapping environments (V_v vs. v_v):

Speaker AW:
/p/: F(2,18) = 22.26; p = .0002
/t/: not collected
/k/: F(2,18) = 1.45; p > .2442
/b/: 0 ms aspiration for all tokens.
/d/: not collected
/g/: 0 ms aspiration for all tokens.

Speaker AM:
/p/: F(2,18) = 69.78; p < .0001
/t/: 0 ms aspiration for all tokens.
/k/: F(2,18) = .10; p > .7541
/b/: 0 ms aspiration for all tokens.
/d/: F(2,18) = 2.95; p > .103
/g/: F(2,18) = 2.26; p > .1499

Speaker RG:
/p/: F(2,18) = 13.30; p < .0018
/t/: F(2,18) = 2.22; p > .1539
/k/: F(2,18) = .6515; p > .4301
/b/: F(2,18) = 3.78; p > .0676
/d/: F(2,18) = 1.00; p > .3306
/g/: F(2,18) = 5.66; p < .0286

Speaker LE:
/p/: F(2,18) = .07; p > .7902
/t/: F(2,18) = 3.78; p > .0677
/k/: F(2,18) = 19.19; p < .0004
/b/: F(2,18) = .01; p > .921
/d/: F(2,18) = 2.23; p > .1529
/g/: F(2,18) = .67; p > .4247