Planning strategies governing selection of kinematic flap/tap variants in North American English

Flap constraints

Donald Derrick and Bryan Gick

University of British Columbia
Categorical behavior in speech is useful for understanding speech motor planning...

... but known cases are rare:

English ‘r’ is famously variable (Delattre and Freeman, 1968; Westbury et al., 1999)
- Tip down bunched
- Tip up bunched
- retroflex
  - ... see also (Tiede et al., 2004; Zhou et al., 2007, 2008)

Ladefoged (1968) distinguished flaps from taps
- Derrick and Gick (2008), distinguish up to four categorically distinct kinematic variations of English flaps/taps:
  - Up-flaps
  - Down-flaps
  - Alveolar taps
  - Postalveolar taps
Examples of 2 kinematic alternatives
- Up-flap vs. down-flap
- Interact w/2 tongue targets: tip up & tip down
  - As in the word Saturday
English speakers may produce some or all of these flap/taps

Hypothesis: this categorical variation is governed by:

- Phonetic constraints
  - Correlate strongly with flap/tap type
  - Show planning within a syllable
  - Have high within and between subject variability
  - ... which helps identify:

- Motor constraints
  - Planning across two syllables/three segments

- Physiological constraints
  - Planning based on tongue speed

- Psychological constraints
  - Strategy shift based on speech errors
Preliminary work suggests that, in single flap/tap sequences, the variants are largely dependent on context:

- English vocalic rhotics (r) can be produced in (at least) two categorically describable ways:
  - With the tongue tip up or down (Delattre and Freeman, 1968) (Above vs. below the alveolar ridge)
    - Tongue tip position from flaps or taps high or low next to vocalic rhotics based on rhotic TT position
  - English vowels (V) are produced with the tongue tip down
    - Tongue tip position from flaps or taps ideally low adjacent to vowels
Hypotheses:

1) VrV ⇒ ‘autumn’
   - alveolar taps.
2) ăr V ⇒ ‘Berta’
   - initial tip-up ‘r’ ⇒ down flaps
   - initial tip-down ‘r’ ⇒ alveolar taps
3) Vră ⇒ ‘Otter’
   - final tip-up ‘r’ ⇒ up-flaps
   - final tip-down ‘r’ ⇒ alveolar taps
4) In a ără ⇒ ‘murder’
   - initial and final tip-up ‘r’ ⇒ post-alveolar taps.
   - initial and final tip-down ‘r’ ⇒ alveolar taps
   - initial tip-down ‘r’, final tip-up ‘r’ ⇒ up-flaps
   - initial tip-up ‘r’, final tip-down ‘r’ ⇒ down-flaps
While regular B-mode (2D) ultrasound displays at 30 frames per second (fps) to video
- Too slow to record flaps or taps properly

Solution M-mode (1D) ultrasound can transfer 3x(60-100) fps of information to video
- Fast enough to track the tongue tip in taps and flaps
Participants
- 24 native American and Canadian English speakers (16 recorded)

Stimuli
- Part of a larger experiment set of 38 sentences repeated 12 times each containing
  - 17 control sentences
  - 9 sentences with 1 flap
  - 10 sentences with 2 flaps
  - 2 sentences with three flaps
Stimuli (cont.)

- Focus on 4 stimuli
  - ‘We have him edify a book’ ⇒ VrV
  - ‘We have Berta beep’ ⇒ ərV
  - ‘We have otter books’ ⇒ Vrə
  - ‘We have him murder a mob’ ⇒ ərə

Recording:
- Ultrasound: Aloka ProSound SSD-5000
- Transducer: 180° EV wand
- Microphone: Sennheiser MKH-416 short shotgun microphone
- PreAmp: M-Audio DMP3 via XLR cable

Digitization:
- ADVC110 Canopus A/D video converter
- iMovie HD (2006)

Analysis:
- Elan 3.5.0-3.8.1
Methods

Setup

(a) Sagittal  
(b) Coronal  
(c) Transverse

**Figure:** Participant Seating

(a) Front  
(b) Back

**Figure:** Canopus Setup
(a) ‘Berta’ ⇒ down-flap. Tongue-surface trajectory moves downward

(b) ‘otter’ ⇒ up-flap. Tongue-surface trajectory moves upward

(c) ‘autumn’ ⇒ alveolar tap. Tongue-surface trajectory moves up and down

(d) ‘murder’ ⇒ postalveolar tap. Tongue-surface trajectory is flat and higher than that of the alveolar tap
Multinomial logistic regression comparing flap type
  - up-flap
  - down-flap
  - alveolar tap
  - postalveolar tap
  vs. tongue tip position both before and after the vowel
  - non-rhotic vowel
  - tip-up rhotic vowel
  - tip-down rhotic vowels

Demonstrates a strong relationship between the two:
Figure: red = alveolar tap, green = up-flap, blue = down-flap, purple = postalveolar tap. (Red/Blue shows diff between tip-down rhotic and non-rhotic vowels)
All four variants exist
  - Some speakers produce flaps but little/no taps
Tip down rhotic and non-rhotic vowels have similar effects
  - ... which differ from tip-up rhotic vowels
Cognitive psychologists:
- Speech planning down to the phoneme level (Levelt, 1989; Dell, 1986)

Speech scientists
- Coarticulated features with:
  - No planning (Öhman, 1966, 1967; Fowler, 1980; Saltzman and Munhall, 1989; Boyce, 1990)
  - Limited planning (Henke, 1966; Whalen, 1990)

Also Munhall et. al. (Munhall et al., 2000) illustrated difficulty in determining which hypotheses is correct
Hypothesis: Flaps will be selected to accommodate end-state comfort Rosenbaum et al. (1992)

- Expect flap types to be largely independent of the initial rhotic
  - Beginning-state
- Expect flap types to be dependent on final rhotic
  - End-state
- Planning based on end-state comfort
Results

Initial rhotic and flap selection: ‘Berta’

(a) Initial tip-down rhotic
(b) Initial tip-up rhotic

Figure: flap/taps in ‘Berta’ based on initial rhotic

No significant difference (GLMM)
Final rhotic and flap selection: ‘otter’

(a) Final tip-down rhotic

(b) Final tip-up rhotic

Figure: flap/taps in ‘otter’ based on final rhotic

Significant difference (generalized linear mixed model: GLMM):
AIC = 262.5, z = 9.93, p = <0.001.
Results distinguish rhotic vowels before and after the flap/tap
  - Flap/tap type influenced by final rhotic type
But...
  - Flap/tap type NOT influenced by initial rhotic type
Results support planning based on end-state comfort
Hypothesis: Flap/tap planning based on end-state comfort extends across two syllables/three phonemes

- initial flap in double flap/tap phrase different due to rhotic two syllables away
  - first flap: ‘We have him edify/audify a book’ vs. ‘We have editor/auditor books’
    - alveolar tap vs. up-flap
  - final flap: ‘We have otter books’ vs. ‘We have editor/auditor books’
    - up-flap vs. postalveolar tap
Results

First flap descriptive

Figure: first flap/tap type comparing ‘Edify/Audify’ vs. ‘Editor/Auditor’
Results

Final flap descriptive

(a) ‘otter’

(b) ‘editor/auditor’

Figure: second flap/tap type comparing ‘otter’ vs. ‘Editor/Auditor’
The results support both hypotheses
  - They are true of 12 of 18 subjects

The results show that a sequence-final rhotic affects initial flap/tap selection
  - 1 syllables/3 segments later
  - Across morpheme boundary
Hypothesis: Tongue speed may affect flap/tap selection
- In single flap/tap sequences
- In the first flap/tap of double flap/tap sequences
- In the second flap/tap of double flap/tap sequences
Faster tongue speed correlates with more taps than flaps
### Results

**Single flap sequences**

Flap kinematics by tongue speed

<table>
<thead>
<tr>
<th>Tongue Speed (ms)</th>
<th>Alveolar Tap</th>
<th>Down Flap</th>
<th>Up Flap</th>
<th>Post-Alveolar Tap</th>
</tr>
</thead>
<tbody>
<tr>
<td>150</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>140</td>
<td>45</td>
<td>35</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td>130</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>120</td>
<td>35</td>
<td>25</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>110</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

(a) Single flap phrases

(b) First flap, double flap phrases

(c) Second flap, double flap phrases
Faster tongue speed correlates with:

- Single flap/tap
  - Less up-flaps
  - More taps
- First of double flap/tap
  - More down-flaps and taps
  - Less up-flaps
- Second of double flap/tap
  - More alveolar taps/up-flaps
  - Less postalveolar taps/down-flaps

Hypothesis supported

Summary: faster tongue speed correlates with more taps vs. flaps
Hypothesis: Speech errors will increase the likelihood of strategy shift

- Speech errors = all speech disfluencies
  - False starts
  - Hesitations
  - Phonetic, phonological, morphological, syntactic errors
  - Wrong words

- Strategy = selection of flap/tap type
- Shift = change in strategy for the same phrase later in the experiment
Results: Speech errors

The results are highly significant [$F(1, 1825) = 11.938, p < 0.001$, $R^2 = 99\%$].

**Figure**: Strategy-shift in relation to most recent speech error/disfluency
Preliminary result supports the hypothesis
... even at very long distances!
There are four categorical kinematic variants of flap/taps in English.

They provide evidence of planning:
- Across one segment
- Across three segments/two syllables

Tongue speed influences strategy.

Speech errors/disfluencies influence strategy:
- Provides evidence of a system seeking equilibrium
- That is, seeking patterns that work consistently
  - Conscious or unconscious?
Special thanks to Aislin Stott for segmenting the underlying acoustic data. This research was funded by a Discovery Grant from the Natural Sciences and Engineering Council of Canada (NSERC) to the second author, and by National Institutes of Health (NIH) Grant DC-02717 to Haskins Laboratories.


