“Peppe[r]oni and g[ə]een peppe[r]s": Evidence of word–internal phonological code switching

Daniel Duncan, New York University
CUNY Phonology Forum, 1–15–15
Introduction

- Primarily morphosyntactic
- Not seen as word–internal (Budzhak–Jones 1998)
Introduction

- Difficulties in studying phonological aspects of code-switching
- Can CS occur in phonology?
  - “Phonological systems cannot be mixed” (MacSwan 2000)
  - Russian–Estonian CS suggests phonological adaptation not necessarily connected to morphological adaptation (Zabrodskaja 2009)
- Supposing it can occur...
  - Phonology is variable
  - Can display L1 interference effects
This Study

- Use mediated staged performances (music) to (mostly) avoid variability/L1 interference issues
- In performance, linguistic features are conscious and pre-mediated (Bell and Gibson 2011)
  - Performer *wants* feature present
  - Performer *has opportunity* to remove unwanted feature by re-recording
Music and Language

- Typically studied with sociolinguistic focus
  - **Style shifting** (Trudgill 1983, Simpson 1999, Gibson and Bell 2012, Duncan 2014, etc.)
  - **Language attitudes** (Beal 2009, Sippola 2014, etc.)
- This study primarily uses music as a corpus of English speech in language contact situation
Road Map

- Introduction
- Problem
- Methodology
- Results
- Discussion
System of a Down

- Armenian–American metal band
- Lead singer is Armenian–English bilingual, attended Armenian–English high school in Los Angeles
- Lyrics in English
- Singer occasionally deploys trilled [r] in songs
System of a Down

Circumventing circuses, lamenting in protest
To visible police, presence sponsored fear
Battalions of riot police with rubber bullet kisses
Baton courtesy, service with a smile
Beyond the Staples Center you can see America
With it's tired poor avenging disgrace
Peaceful loving youth against the brutality
Of plastic existence

(“Deer Dance”, 2001)
System of a Down

- Use of [r] appears to be intentional stylistic choice
- Deployed in same word in both recorded and live formats

The call of the righteous man (“War?”, 1998)
Armenian Phonology

- Armenian has /r,ɾ/ in consonant inventory

- Rules applying to rhotic distribution (Vaux 1998)
- Final Devoicing
  - Some word–final continuants are devoiced
  - Often approaches [ʃ] in some dialects
- Word–Initial Consonant Clusters
  - Epenthetic [ə] between underlying consonants
Armenian Phonology

- Surface phonotactic distribution:
  - Word–final R restricted
  - Word–initial CR clusters restricted
Problem

- Use of [r] appears to be Armenian influence
- Is this borrowing or code-switching?
Road Map

- Introduction
- Problem
- Methodology
- Results
- Discussion
Variationist Methodology

  - Borrowing: [r] use matches English /ɹ/
  - Code-switching: [r] use matches Armenian /ɾ/

- Focus on two conflict sites (Poplack and Meechan 1998)
  - Word-initial consonant clusters
  - Word-final appearance
Variationist Methodology

- Corpus of 67 songs (five albums) on Spotify
- Impressionistically collected all tokens with r (~3400)
- Tokens labeled as trill or English
- Additionally coded for album, word position (initial/medial/final), CC-cluster
- Exclusions:
  - [r]—mostly in non-English words
  - /θɹ/—Armenian lacks /θ/
Road Map

- Introduction
- Problem
- Methodology
- Results
- Discussion
Results

- Use of [r] actually quite rare: 32 of 3375 tokens
- Raw frequencies suggest difference in distribution between [r,ɹ]
## Results

<table>
<thead>
<tr>
<th></th>
<th>[i]</th>
<th></th>
<th>[r]</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Initial</td>
<td>891</td>
<td>26.65%</td>
<td>16</td>
<td>50.00%</td>
</tr>
<tr>
<td>Medial</td>
<td>761</td>
<td>22.76%</td>
<td>12</td>
<td>37.50%</td>
</tr>
<tr>
<td>Final</td>
<td>1691</td>
<td>50.58%</td>
<td>4</td>
<td>12.50%</td>
</tr>
<tr>
<td>Total</td>
<td>3343</td>
<td>100%</td>
<td>32</td>
<td>100%</td>
</tr>
</tbody>
</table>

Raw frequencies of tokens by word position
## Results

<table>
<thead>
<tr>
<th></th>
<th>[ɪ]</th>
<th>[r]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consonant Cluster</strong></td>
<td>1371</td>
<td>10</td>
</tr>
<tr>
<td><strong>No Cluster</strong></td>
<td>1972</td>
<td>22</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>3343</td>
<td>32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>[ɪ]</th>
<th>[r]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consonant Cluster</strong></td>
<td>41.01%</td>
<td>31.25%</td>
</tr>
<tr>
<td><strong>No Cluster</strong></td>
<td>58.99%</td>
<td>68.75%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Raw frequencies of tokens by CC–cluster presence
Results

- Fixed-effects logistic regression in Rbrul (Johnson 2009)
- Each factor is significant predictor ($p < .05$)
- Use of [r] diminished across career output
# Results

<table>
<thead>
<tr>
<th>Input</th>
<th>Centered Weight</th>
<th>% [r]</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Album</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypnotize</td>
<td>.705</td>
<td>0.38%</td>
<td>518</td>
</tr>
<tr>
<td>Mezmerize</td>
<td>.577</td>
<td>0.63%</td>
<td>640</td>
</tr>
<tr>
<td>Steal This Album</td>
<td>.539</td>
<td>0.74%</td>
<td>808</td>
</tr>
<tr>
<td>System of a Down</td>
<td>.379</td>
<td>1.26%</td>
<td>635</td>
</tr>
<tr>
<td>Toxicity</td>
<td>.301</td>
<td>1.55%</td>
<td>774</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>.404</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total N=3375

Values <.5 favor [r], values >.5 disfavor [r]
# Results

<table>
<thead>
<tr>
<th>Album</th>
<th>Centered Weight</th>
<th>% [r]</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>#5</td>
<td>0.705</td>
<td>0.38%</td>
<td>518</td>
</tr>
<tr>
<td>#4</td>
<td>0.577</td>
<td>0.63%</td>
<td>640</td>
</tr>
<tr>
<td>#3</td>
<td>0.539</td>
<td>0.74%</td>
<td>808</td>
</tr>
<tr>
<td>#1</td>
<td>0.379</td>
<td>1.26%</td>
<td>635</td>
</tr>
<tr>
<td>#2</td>
<td>0.301</td>
<td>1.55%</td>
<td>774</td>
</tr>
</tbody>
</table>

**Range** | 0.404

Values <.5 favor [r], values >.5 disfavor [r]
Results

- Fixed-effects logistic regression in Rbrul (Johnson 2009)
- Each factor is significant predictor (p<.05)
- Use of [r] diminished across career output
- Contexts disfavoring use of [r]
  - Word–final position
  - Presence in CC–cluster
## Results

<table>
<thead>
<tr>
<th>Position</th>
<th>Centered Weight</th>
<th>% [r]</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final</td>
<td>.832</td>
<td>0.24%</td>
<td>1695</td>
</tr>
<tr>
<td>Medial</td>
<td>.325</td>
<td>1.55%</td>
<td>773</td>
</tr>
<tr>
<td>Initial</td>
<td>.295</td>
<td>1.76%</td>
<td>907</td>
</tr>
<tr>
<td>Range</td>
<td>.537</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cluster</th>
<th>% [r]</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster</td>
<td>0.72%</td>
<td>1381</td>
</tr>
<tr>
<td>No cluster</td>
<td>1.10%</td>
<td>1994</td>
</tr>
<tr>
<td>Range</td>
<td>.282</td>
<td></td>
</tr>
</tbody>
</table>

Input: .994  
Total N=3375

Values <.5 favor [r], values >.5 disfavor [r]
Results

- Cross-tabulation highlights difference between word-initial clusters vs. non-clusters

<table>
<thead>
<tr>
<th></th>
<th>Word-Initial</th>
<th></th>
<th>Word-Medial</th>
<th></th>
<th>Word-Final</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>No consonant</td>
<td>12</td>
<td>3.45%</td>
<td>7</td>
<td>1.99%</td>
<td>3</td>
<td>0.23%</td>
<td>22</td>
<td>1.10%</td>
</tr>
<tr>
<td>cluster</td>
<td>349</td>
<td></td>
<td>352</td>
<td></td>
<td>1293</td>
<td></td>
<td>1994</td>
<td></td>
</tr>
<tr>
<td>Consonant</td>
<td>4</td>
<td>0.72%</td>
<td>5</td>
<td>1.12%</td>
<td>1</td>
<td>0.25%</td>
<td>10</td>
<td>0.72%</td>
</tr>
<tr>
<td>cluster</td>
<td>558</td>
<td></td>
<td>421</td>
<td></td>
<td>402</td>
<td></td>
<td>1381</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>1.76%</td>
<td>12</td>
<td>1.55%</td>
<td>4</td>
<td>0.24%</td>
<td>32</td>
<td>0.95%</td>
</tr>
<tr>
<td></td>
<td>907</td>
<td></td>
<td>773</td>
<td></td>
<td>1695</td>
<td></td>
<td>3375</td>
<td></td>
</tr>
</tbody>
</table>

Cross-tabulated results by word position and cluster presence
Results

- Chi square test: Difference between clusters and non-clusters only significant word-initially
  - Initial: $p=0.0035$
  - Medial: $p=0.3964$
  - Final: $p=1.0000$
Road Map

- Introduction
- Problem
- Methodology
- Results
- Discussion
### Comparison of Armenian–English in conflict sites

<table>
<thead>
<tr>
<th></th>
<th>Armenian</th>
<th>English</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word–initial clusters</td>
<td>/r/ restricted</td>
<td>/ɹ/ allowed</td>
<td></td>
</tr>
<tr>
<td>Word–final appearance</td>
<td>/r/ restricted</td>
<td>/ɹ/ allowed</td>
<td></td>
</tr>
</tbody>
</table>
## Discussion

<table>
<thead>
<tr>
<th></th>
<th>Armenian</th>
<th>English</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word–initial clusters</td>
<td>/r/ restricted</td>
<td>/ɹ/ allowed</td>
<td>[r] disfavored</td>
</tr>
<tr>
<td>Word–final appearance</td>
<td>/r/ restricted</td>
<td>/ɹ/ allowed</td>
<td>[r] disfavored</td>
</tr>
</tbody>
</table>

Comparison of Armenian–English in conflict sites
Discussion

- Distribution of [r] matches Armenian restrictions, not English, in conflict sites.
- Variationist approach suggests this represents use of Armenian phonology:
  - Would have mirrored English distribution if borrowed allophone.
  - Lexicon, other phonological features are English, however.
- Appears to be evidence of word–internal code–switching.
Future Directions

- Performative speech appears to be suitable for this research
- Current study one feature, one individual
  - Further work on other bilinguals, more features
- Social aspect of performance needs consideration
Bibliography


Discography

-- 2002. *Steal This Album!*
What’s /θɹ/ Doing?

- Only case of English words with [ɾ]
- Token *through* appeared in very repetitive lyrics
- [ɾ] occurred at quite different rate from other CC–clusters (p<<.0001)

<table>
<thead>
<tr>
<th></th>
<th>/θɹ/</th>
<th>Other CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ɾ]</td>
<td>14</td>
<td>25.9%</td>
</tr>
<tr>
<td>[ɾ]</td>
<td>12</td>
<td>22.2%</td>
</tr>
<tr>
<td>[ɹ]</td>
<td>28</td>
<td>51.9%</td>
</tr>
</tbody>
</table>

Distribution of /θɹ/ tokens vs. other clusters