Prosodic and Segmental Conversion Between Speakers of Different Dialects

Jelena Krivokapić
Linguistics Department
Yale University
jelena.krivokapic@yale.edu

My thanks to Jessica Hsieh and Tine Mooshammer.
Goals

- Nature of speaker convergence
  - How are \textit{prosodic} characteristics affected?
    - Examining whether convergence occurs in suprasegmental properties (\textit{pitch accent} and \textit{prosodic boundary} placement)
  - What is the effect of interaction between speakers of different dialects?
    - Examining whether convergence occurs across dialects (British and American English)
Speaker Convergence

- Speakers’ productions change and become more similar to their co-speakers’ and to their language environment, converging in
  - syntactic structure
    e.g., Bock 1986, Pickering and Ferreira 2008
  - lexical choice
    e.g., Garrod and Doherty 1994
  - phonetic properties
Convergence as Gestural Drift

- Convergence can be explained by speakers’ inclination towards imitation

- Perception of articulatory gestures drives production, leading to automatic gestural attunement
  - Gestural drift: “perceptually guided changes in speech production”
    Sancier & Fowler 1997

- Can form the basis for continuous language learning and accent change
Prosodic Convergence Across Dialects (Ní Chiosáin (2007))

- Ní Chiosáin (2007) examines interaction between two dialects of Irish, using the synchronous speech paradigm.
- Investigated variables: lexical stress, vowel duration, lenition
- Only small effects of convergence found, mainly for lexical stress
In a first analysis of the data presented here, interaction leads to convergence for both segmental and prosodic properties, although only for some subjects (Krivokapić 2010).

- Two British and one American subject show convergence for vowel quality.
- Two British and one American subject showed convergence for stress.
- Intonational contours show indication of reduced variability.
Convergence on occurrence of **prosodic boundaries** and **pitch accents** examined

- Both expected to converge, similar to segmental convergence
  - Like constriction gestures, prosodic boundaries and tones can be modeled as gestures
  Byrd & Saltzman 2003, Gao 2008

- Prosodic boundaries are largely determined by syntactic structure. They are expected to largely converge to a default
  see also Zvonik and Cummins 2002, 2003

- Pitch accents are less predictable. They are expected to converge, but to a lesser extent than boundaries.
**Synchronous Speech Paradigm**


- Minimizes individual, non-linguistic variation without inducing artificial temporal properties

- Dyads read a text simultaneously
  - seated facing each other and recorded on stereo channels with head-mounted microphones

- Captures in a unique way speakers’ shared knowledge of linguistic timing
  Cummins 2002

- Reduces variability in pause duration and F0 contour

- Used here as a way to induce convergence
Experiment Design

- Acoustic study

- A short story containing 10 test words differing between the dialects in vowel quality and 4 words differing in stress placement
  
  Test words from Wells 1982, Berg 1999, story adapted from Honorof, McCullough & Somerville 2000

- Recordings on separate days for solo and synchronous condition
  
  - solo condition always first

- In each session, subjects read 6 repetitions of the story, with 84 filler sentences between each repetition.
The nurse took a bath when she woke up. Then she put on a plain yellow dress and a fleece jacket, picked up her goose Nico, and headed north to work. She worked with her father in a company that produced cloth for cleaning crystal. But a strange thing happened when she walked out of the door: She saw a goat standing near the big garage where she kept her coffee and tools. That reminded her of a story where a dictator had a square hat and always talked about caffeine at press conferences. She was a young adult when she first heard the story, and her favorite meal was tuna with parsley.
Stimuli

- For each speaker **pitch accent** and **boundary placement** analyzed in solo and synchronous condition
  - In each reading of the story, 78 words that can get a pitch accent and occur only once in the story
  - In each reading of the story, 112 words that can be followed by a boundary
Subjects

- 8 speakers (4 dyads)
  - 4 native speakers of American English
  - 4 native speakers of British English

Length of stay in the US:
- recently arrived: 3 months, 1 year
- long time residents: 5 years, 19 years
Data Labeling

- Perceptual prosodic data analysis
  - boundary occurrence (but not strength) in solo and synchronous condition after each word
  - pitch accent occurrence (but not type) for each word

- Two labelers, labeling 100% and 20% of the data
  - Labeler agreement: boundary occurrence 91%, pitch accent occurrence 84.11%
Classification of Changes

Comparison of number of pitch accents and boundary occurrences for each word, across repetitions, in synchronous and solo condition.

- **Convergence**: one speaker, who had consistent production in solo speech (all 6 reps same), changed in the direction of the co-speaker, who also had consistent productions in solo (at least 5 productions same) and synchronous speech (all 6 reps same).

- **Opposite**: Same conditions as above, but speaker changes production in opposite direction from co-speaker

- **Consistent**: Speaker decreases variability in own production from solo to synchronous speech

- **Other**: Speaker changes production, but pattern unclear
Results: Pitch Accents

- Excluded words for which both speakers of a dyad had the same number of pitch accents in solo and synchronous condition.
  - From 312 (78 possible pitch accents x 4 dyads), 107 words remained in the analysis (34.29%)
  - Dyad 1: 26, Dyad 2: 24, Dyad 3: 28, and Dyad 4: 29

=> Speakers have fairly consistent pitch accent placement
## Results: Pitch Accents

<table>
<thead>
<tr>
<th></th>
<th>Dyad 1 26 words</th>
<th>Dyad 2 24 words</th>
<th>Dyad 3 28 words</th>
<th>Dyad 4 29 words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Brit</td>
<td>Amer</td>
<td>Brit</td>
<td>Amer</td>
</tr>
<tr>
<td>Converge</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Opposite</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Consist.</td>
<td>5</td>
<td>11</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>5</td>
<td>14</td>
<td>11</td>
</tr>
</tbody>
</table>

★ recently arrived British speakers

<table>
<thead>
<tr>
<th></th>
<th>British pooled (107 tokens)</th>
<th>American pooled (107 tokens)</th>
<th>All speakers pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nr.</td>
<td>%</td>
<td>Nr.</td>
</tr>
<tr>
<td>Converge</td>
<td>11</td>
<td>10.28</td>
<td>4</td>
</tr>
<tr>
<td>Opposite</td>
<td>2</td>
<td>1.87</td>
<td>0</td>
</tr>
<tr>
<td>Consist.</td>
<td>25</td>
<td>23.36</td>
<td>29</td>
</tr>
<tr>
<td>Other</td>
<td>34</td>
<td>31.77</td>
<td>26</td>
</tr>
</tbody>
</table>
Results: Prosodic Boundaries

- Excluded words for which both speakers of a dyad had the same number of boundaries in solo and synchronous condition.

- From 448 (112 possible boundaries x 4 dyads) 70 words (15.62%) remained in the analysis
  - Dyad 1: 18, Dyad 2: 13, Dyad 3: 21, Dyad 4: 18

=> Speakers have very consistent boundary placement
### Results: Prosodic Boundaries

<table>
<thead>
<tr>
<th></th>
<th>Dyad 1 18 words</th>
<th>Dyad 2 13 words</th>
<th>Dyad 3 21 words</th>
<th>Dyad 4 18 words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Brit</td>
<td>Amer</td>
<td>Brit</td>
<td>Amer</td>
</tr>
<tr>
<td>Converge</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Opposite</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Consist.</td>
<td>2</td>
<td>3</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
<td>10</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>

- *recently arrived British speakers*

<table>
<thead>
<tr>
<th></th>
<th>British pooled (70 words)</th>
<th>American pooled (70 words)</th>
<th>All speakers pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nr.</td>
<td>%</td>
<td>Nr.</td>
</tr>
<tr>
<td>Converge</td>
<td>1</td>
<td>1.43</td>
<td>2</td>
</tr>
<tr>
<td>Opposite</td>
<td>1</td>
<td>1.43</td>
<td>3</td>
</tr>
<tr>
<td>Consist.</td>
<td>17</td>
<td>24.28</td>
<td>12</td>
</tr>
<tr>
<td>Other</td>
<td>17</td>
<td>24.28</td>
<td>28</td>
</tr>
</tbody>
</table>
Summary

- Evidence of convergence of pitch accent
- No evidence of convergence of prosodic boundaries

Possible explanations:
- default prosodic phrasing largely produced in solo reading already
- future work: examine whether convergence is observed in the temporal properties of prosodic boundaries
- repetitive nature of the task leads to a more consistent production of phrasing
- Strong tendency of speakers to become more stable in their placement of both pitch accent and boundary
Conclusions

- Speaker interaction can lead to prosodic convergence in pitch accent placement
  - Similar to segmental properties
- No convergence in prosodic boundary placement
- Further evidence for convergence across dialects
References


Summary

- Evidence of convergence of pitch accent
- No evidence of convergence of prosodic boundaries

Possible explanations:
- default prosodic phrasing has largely been produced in solo reading already
  - future work: examine whether convergence is observed in the temporal properties of prosodic boundaries
- repetitive nature of the task leads to a more consistent production of phrasing

- Strong tendency of speakers to become more stable in their placement of both pitch accent and boundary