Sound change propagation: the relation between perception and production in individual language users

Patrice Speeter Beddor
Andries W. Coetzee
- Variable perception
- Variable perception in relation to production—and sound change?
- How we are approaching these issues in our lab
Perception is malleable and dynamic

- is continuously retuned
  - context-dependent (compensation for coarticulation)
  - cue-dependent (trading relations)
  - rate-specific (Repp et al. 1978)
  - talker-specific (Mann & Repp 1980)
  - new experiences (Kraljic et al. 2008)

- evolves in real time as input acoustic signal unfolds

- varies across listeners
  - different listeners make different linguistic decisions based on the same acoustic input
Malleable, dynamic perception contributes to sound change

- listener misapprehension
  (Ohala 1981, and subsequently many others)

- experience with new phonetic patterns

- even *stable* patterns of variation can lead different listeners to assign different perceptual weights to properties of input
  (Beddor 2009, Yu 2010, Beddor et al. 2013, Garrett & Johnson 2013)
Individual listener variation

Beddor, McGowan, Boland, Coetzee, & Brasher 2013

- Used eye-tracking methods to track the time course of English-speaking listeners' use of Ŵ as information for an upcoming N.

- Many listeners begin to fixate a CVNC image rather than a CVC image (e.g., *bend* rather than *bed*) nearly as soon as Ŵ begins.
  - Some, but not all, of these listeners continue to fixate the CVNC image even when there is no acoustic N (e.g. [bẽd]).

- Other listeners: systematically delay looks to CVNC until hear N; Ŵ not very useful information.
Direct manifestation would be a systematic relation between a language user's percept and that individual's productions: e.g., perhaps

- Information heavily weighted in perception is especially prominent in that language user's productions.
- Shift in perceptual category boundaries due to novel experiences leads to shift in production norms for that language user.
- Incorrect percept (misparsing) is replicated in production in listener-turned-speaker (Ohala).
What is the nature of the perception–production link?

Should not expect a precise mirror

- Expect greater malleability in perception than in production
- Flexibility in perception is essential to comprehension; however, non-accommodation presumably often does not preclude intelligibility (Pardo 2012)
- Garrett & Johnson (2013): propose a dual-representation model in which perception and production are based on different sets of exemplars
What is the nature of the perception–production link?

Should not expect a precise mirror

- And we don't find one!
## Perception-production relation?

<table>
<thead>
<tr>
<th>Perception parallels production</th>
<th>Perception ≠ production</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Grosvold &amp; Corina (2012)</td>
</tr>
<tr>
<td>Harrington et al. (2008)</td>
<td>Kraljic et al. (2008)</td>
</tr>
<tr>
<td>Kleber et al. (2011)</td>
<td></td>
</tr>
</tbody>
</table>
### Perception-Production Relation?

<table>
<thead>
<tr>
<th>Perception parallels production</th>
<th>Perception ≠ production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perkell et al. (2004a, 2004b)</td>
<td>Participants who PRODUCE larger differences between members of a native contrast (e.g., /u/-/ʊ/) more accurately DISCRIMINATE that contrast.</td>
</tr>
<tr>
<td>Ghosh et al. (2010)</td>
<td></td>
</tr>
<tr>
<td>Schwartz et al. (2012)</td>
<td></td>
</tr>
<tr>
<td>Harrington et al. (2008)</td>
<td></td>
</tr>
<tr>
<td>Kleber et al. (2011)</td>
<td></td>
</tr>
</tbody>
</table>
### Perception-parallels production

<table>
<thead>
<tr>
<th>Perception parallels production</th>
<th>Perception ≠ production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schwartz et al. (2012)</td>
<td></td>
</tr>
<tr>
<td>Harrington et al. (2008)</td>
<td></td>
</tr>
<tr>
<td>Kleber et al. (2011)</td>
<td></td>
</tr>
</tbody>
</table>

Speakers' PRODUCED F1 distances between vowels that differ in height correlate with PERCEIVED F1 distances.
### Perception—production relation?

<table>
<thead>
<tr>
<th>Perception parallels production</th>
<th>Perception ≠ production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harrington et al. (2008)</td>
<td></td>
</tr>
<tr>
<td>Kleber et al. (2011)</td>
<td></td>
</tr>
</tbody>
</table>

Younger speakers both PRODUCE /u/ with fewer coarticulatory influences and PERCEPTUALLY adjust less for these influences than do older speakers.
<table>
<thead>
<tr>
<th>Perception parallels production</th>
<th>Perception ≠ production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants who produce more extensive coarticulation do NOT more accurately perceive differences between coarticulated variants (Grosvald) nor do they show larger perceptual boundary shifts (Kataoka).</td>
<td></td>
</tr>
<tr>
<td>Kataoka (2011)</td>
<td></td>
</tr>
<tr>
<td>Kraljic et al. (2008)</td>
<td></td>
</tr>
</tbody>
</table>
### Perception-Production Relation?

<table>
<thead>
<tr>
<th>Perception parallels production</th>
<th>Perception ≠ production</th>
</tr>
</thead>
</table>

Perceptual learning of new variants did not result in corresponding shift in speakers' spontaneous productions.
In our lab: cue weighting

- Do language users who *produce* more innovative variants also weight the innovative property more heavily in *perception*?

- Current projects:
  - Ongoing sound change in Afrikaans  
    (with Daan Wissing)
  - Stable patterns of coarticulatory variation in English  
    (with Kevin McGowan, Andries Coetzee, Julie Boland, Anthony Brasher, Cameron Rule)
Afrikaans devoicing / tonogenesis (Coetzee, Beddor, Wissing)

- Afrikaans traditionally described as contrasting prevoiced and voiceless unaspirated stops

- Examples:
  - *bak* [bak] ‘bowl’
  - *pak* [pak] ‘pack’
Afrikaans devoicing

- Afrikaans traditionally described as contrasting prevoiced and voiceless unaspirated stops
- But "voiced" stops may lack prevoicing
- /voiced/-/voiceless/ difference can be heard in post-stop f0

*bak* [pak] ‘bowl’

*pak* [pak] ‘pack’
Afrikaans devoicing

- Afrikaans traditionally described as contrasting prevoiced and voiceless unaspirated stops
- But "voiced" stops may lack prevoicing
- /voiced/-/voiceless/ difference can be heard in post-stop f0

**Is f0 becoming the primary information for (initial) voicing contrasts?**
(exaggeration of low-level f0 perturbations after voiceless vs. voiced consonants)

**Are language users who produce less voicing (larger f0 differences?)**
also perceptually less sensitive to voicing information?
Production

- Afrikaans speakers (all female)
  - 13 older (12 over 50 years; one: 40 years)
  - 10 younger (under 25)

- Stimuli
  - Labial: 15 /b/-initial words: bas, baal, bied, ...
  - 15 /p/-initial words: pas, paal, Piet, ...
  - Alveolar: 15 /d/-initial words: das, dak, doer, ...
  - 15 /t/-initial words: tas, tak, toer, ...

  Fillers that begin with other consonants.

- Measures: VOT and f0 across post-stop vowel
Production

- Predictions
  - Younger speakers will devoice more.
  - Contrast will be maintained in f0.
  - Exaggerated f0 contrast for devoiced /b d/?
Production

- Predictions
  - Younger speakers will devoice more.
  - Contrast will be maintained in f0.
  - Exaggerated f0 contrast for devoiced /b d/?
VOT Production

Afrikaans: production

Token count

VOT (ms)

Voiced  Voiceless

/p/
/b/
/t/
/d/

Less -200 -180 -160 -140 -120 -100 -80 -60 -40 -20 0 20 40 60 80 More

Less -200 -180 -160 -140 -120 -100 -80 -60 -40 -20 0 20 40 60 80 More

0 20 40 60 80 100 120 140 160 180 200 Token count
VOT Production

% tokens produced as voiced

Younger speakers devoice more than older.
Afrikaans: production

**f0 Production**

<table>
<thead>
<tr>
<th>Older</th>
<th>Labial</th>
<th>Alveolar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td></td>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

Normalized vowel duration

*f0 depends on phonological voicing.*
### Afrikaans: production

#### f0 Production

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Alveolar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Older</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td>Younger</td>
<td><img src="image3" alt="Graph" /></td>
<td><img src="image4" alt="Graph" /></td>
</tr>
</tbody>
</table>

- /voiceless/ - yellow line
- voiced - blue line
- devoiced - dotted line

Normalized vowel duration

- f0 is not age-dependent.
Production Summary

- f0 is stable
  - Depends on phonological voicing
  - Voicers (mainly older speakers) and devoicers alike produce large f0 distinctions.

- VOT is variable
  - Younger speakers devoice more than do older speakers.
Perception

What happens when f0 and voicing are orthogonally varied and presented to the same language users?

If perceptual weights parallel production, predict that:

- In absence of voicing, all participants will use f0.
- When f0 and voicing conflict:
  - Voicers (mostly older participants) will rely more on voicing.
Perception

Stimuli

Labial: bas ~ pas /bas/ ~ /pas/
Alveolar: doer ~ toer /du:r/ ~ /tu:r/

- Seven f0 steps = equally spaced between f0 profile of original voiced and voiceless endpoint.
Afrikaans: perception

- Three voicing levels:
  - *Full* = average duration and intensity of prevoicing
  - *Reduced* = 1.5 standard deviation below average
  - *None* = voiceless

Do you hear: bas doer pas toer
Results

Labial

Alveolar

Younger

Older

% voiced response

Low F0 2 3 4 5 6 High F0

Full Reduced None

Low F0 2 3 4 5 6 High F0

*
Results: f0

listeners use f0 in absence of voicing.
Use of f0 is NOT age-dependent.
Results: f0

% voiced response on the “No Voicing” continua

Use of f0 is NOT age-dependent.
Results: voicing

Use of voicing IS age-dependent.
Results: voicing

Use of voicing IS age-dependent.

% voiced response on the “Full/Reduced Voicing” continua
Perception Summary

- Use of f0 is stable
  - No difference between older and younger listeners in perceptual use of f0 when no voicing present.

- Use of voicing is variable
  - When voicing present, older listeners weight voicing more heavily than do younger listeners.

- Some listeners: voicing overrides f0
Individual Listeners

**Labial**

- Heavy weight for voicing

**Alveolar**

- Lighter weight for voicing
Afrikaans: perception

Individual Listeners

Labial

Alveolar

Similar weights for f0
Afrikaans: perception

Individual Listeners

**Labial**

**Alveolar**

No listener required voicing to hear /b d/.

Similar weights for f0
Individual Listeners

Does perceptual weighting correlate with that language user's productions?

Heavy weight for voicing

Lighter weight for voicing
% voiced response to Full/Reduced stimuli vs. % /b/ and /d/ produced with voicing

$r^2 = 0.12, p = 0.10$

Afrikaans: perception-production relation for voicing
The more users produce voicing, the more heavily they weight voicing in perception.

Afrikaans: perception-production relation for voicing
o Have missed expected early stage when f0 was variable

o Now: f0 stable in both perception and production

o **Voicing is the variable property:** the language users who produce more voicing weight voicing more heavily in perception.
• Do language users who *produce* more "innovative" variants also weight the innovative property more heavily in *perception*?

  - Ongoing tonogenesis in Afrikaans:

    ➢ Innovative property: f0 contour

    All speakers: produce large f0 differences and use f0 in perception.

    More accurate to say that the users who produce more conservative (voiced) variants weight the more conservative property more heavily in perception.
Do language users who *produce* more "innovative" variants also weight the innovative property more heavily in *perception*?

- Time course of perception and production of coarticulation: Stay tuned ...
Collaborators

Julie Boland
Anthony Brasher
Cameron Rule
Kevin McGowan
Daan Wissing
Beddor et al. (2013) The time course of perception of coarticulation. *JASA* 133, 2350-2366.
Harrington et al. (2008) Compensation for coarticulation, /u/-fronting, and sound change in SSB. *JASA* 123, 2825-2835.
Perkell et al. (2004a) The distinctness of speakers’ productions of vowel contrasts is related to their discrimination of the contrasts. *JASA* 116, 2338-2344.