¿Quién lo di[s]e? Effects of talker dialect on lexical decision involving a merged phoneme

Duna Gylfadottir
University of Pennsylvania
gug@sas.upenn.edu

SVALP
April 1st, 2016
What is the role of talker voice attributes in linguistic perception?

Specifically, how does talker accent affect processing?
Listeners perceive, retain, and use talker-specific information.

- Improved perception performance for single talker over multiple talker conditions (Sommers et al. 1997) and for familiar voices over unfamiliar ones (Nygaard et al. 1994).

- Exposure periods decrease the difficulty of processing a non-native accent (Clarke and Garrett 2004).

- Exposure to a talker with a shifted set of vowels aids in subsequent comprehension of the talker’s speech (Maye et al. 2008).
Speakers accumulate implicit and explicit knowledge about dialects.

- Speech in an unfamiliar dialect is harder to process.
- Speakers can identify dialects by hearing them (Clopper and Pisoni 2004).
Listeners extract and use information about specific talkers. 
Listeners know about accents and recognize them.

“Listeners can explicitly categorize unfamiliar talkers by regional dialect with above-chance performance under ideal listening conditions. However, the extent to which this affects speech processing is largely unknown.” (Clopper and Bradlow 2008)
Claim: listeners expect features that co-occur in a variety to co-occur in talkers they identify as speaking that variety

- Exposure to a talker with a recognizable accent should generate *expectations* about what that talkers future utterances will sound like.

  - Should not be limited to sounds that they have already heard from that listener: they should extend to other sounds.

→ We can test this by *violating* this expectation.
Research Questions

- What is the lexical status of a word produced with a phoneme consistent with some dialect that is not the listener’s?
  - E.g., is [flɛndə] for slender different from [slɛndə] for slender? (Sumner and Samuel 2009; Sumner and Kataoka 2013)

- Is it easier to process a phoneme substitution when the talker’s accent matches the substitution?
  - E.g., [slɛndə] for slender produced by a talker with BrE features vs. GA features

- Do all listeners use dialect information the same way?
What is it that makes a variant unexpected or salient? (Sumner, SVALP 2016)

- Sociolinguistic (intra-dialect) variants and frame features (Vaughn and Kendall; Tamminga, SVALP 2016)
- Syntactic variants and phonological context (Hanulíková et al. (2012))
- Variants and other kinds of contextual information (Babel, SVALP 2016, Hay and Drager 2010)
## The variable

<table>
<thead>
<tr>
<th>Distinction</th>
</tr>
</thead>
</table>
| dice [di.θe] ‘say’  
base [ba.se] ‘base’  

<table>
<thead>
<tr>
<th>Seseo</th>
</tr>
</thead>
</table>
| dice [di.se] ‘say’  
base [ba.se] ‘base’  

<table>
<thead>
<tr>
<th>Ceceo</th>
</tr>
</thead>
</table>
| dice [di.θe] ‘say’  
base [ba.θe] ‘base’  

Figure 1: Status of /s/ and /θ/ merger in Spain
Task

- Continuous auditory lexical decision task
- Two talker conditions
  - Mexican and Castilian
  - Talkers presented in two blocks
  - Introduced by a short passage of speech with no environments where [s] or [θ] could appear
- 10 Castilian (unmerged) participants
- All /s/- and /θ/-words produced with [s], for both the Mexican and the Castilian conditions
### Table: Stimulus Expectation and Production

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Castilian Talker</th>
<th>Mexican Talker</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expectation</strong></td>
<td><code>θ</code></td>
<td><code>s</code></td>
</tr>
<tr>
<td><strong>Production</strong></td>
<td><code>s</code></td>
<td><code>s</code></td>
</tr>
</tbody>
</table>

*Note: The table compares the expected and produced stimuli for two talkers, one from Castilian and one from Mexican dialects.*
Test stimuli

- Words containing /s/ or /θ/, not minimal pairs
- Frequency balanced /s/ words and /θ/ words
- Proportional number of /s/ pseudowords
- Overall 60:40 word:nonword
- Overall < 20% [s]
Experimental design

- Two blocks, one for each talker
- Same stimuli repeated in the second block in a different order
- Block order varied across subject
Passage Block Passage Block

Castilian talker

Mexican talker
Figure 2: Reaction time for different categories of test stimuli by talker condition.
Results

- All subjects identified the Castilian voice as Spanish and the Mexican voice as not Spanish.
- /s/-words faster than /θ/-words overall ($\beta = 0.166$, $p < .01$)
- /θ/-words accepted faster when the talker was Mexican ($\beta = -0.307$, $p < .01$)
Figure 3: Reaction time for different categories of test stimuli by talker condition
Results

- /s/-words faster than /θ/-words overall ($\beta=0.166$, $p<.01$)
- /θ/-words accepted faster when the talker was Mexican ($\beta=-0.307$, $p<.01$)
- Marginally faster to respond to Castilian talker overall ($\beta=0.037$, $p<.10$)
- A number of expected effects (frequency, syllable structure, rejection delay)
- No improvement over the course of a block (no effect of or interactions with item order)
- No effect or interaction of block order
**Figure 4:** Reaction time for different categories of test stimuli by talker condition.
Figure 5: Percent acceptance of /θ/-words by talker and subject, separated by response strategy.
Figure 6: Reaction time to /θ/-words by talker and subject, separated by response strategy.
Figure 7: Lexical decisions for /θ/-words by each subject across the experiment.
What is the lexical status of a word produced with a phoneme consistent with some dialect but not the listener’s?

NOT CLEAR

- /s/-words significantly faster than /θ/-words overall.
- They do not appear to be pseudowords, and they are also not processed like words.
What is the lexical status of a word produced with a phoneme consistent with some dialect but not the listener’s? NOT CLEAR

Is it easier to process a phoneme substitution when the talker’s accent matches the substitution? YES

- Participants responded to /θ/-type words significantly faster in the Mexican talker condition.
Discussion

- What is the lexical status of a word produced with a phoneme consistent with some dialect but not the listener‘s? NOT CLEAR
- Is it easier to process a phoneme substitution when the talker’s accent matches the substitution? YES
- Do listeners use dialect information the same way? NO
  - Some subjects rejected or accepted /θ/-type words in both conditions while others responded based on talker accent.
  - This did not straightforwardly relate to response times.
What is the lexical status of a word produced with a phoneme consistent with some dialect but not the listener’s? NOT CLEAR

Is it easier to process a phoneme substitution when the talker’s accent matches the substitution? YES

Do listeners use dialect information the same way? NO
Discussion

- Speakers seem to be responding to the incongruency, but not all in the same way
  - Tapping into prescriptive judgments for some participants?
  - What would we expect for more and less salient/stigmatized features?
  - What is the role of cross-dialect experience? (cf., Sumner and Samuel 2009)
    - All things equal, do some individuals use talker accent information more or more effectively than others?
- What is the time course of the use of this information?
  - Is this kind of effect present at earlier stages of processing?
  - What is the effect of this kind of incongruency on memory?
- Where does this fit in the effect of the style of frame sentence on judgments and processing of stigmatized variants, and the effects of face attributes, explicit dialect labels, etc. on perception?
Conclusions

- Study shows that dialect information present in the signal affects lexical decision
  - Without being given explicit dialect labels.
  - Without exposure to the relevant contrast.
- Indicates integration of pre-existing dialect information into responses to novel talkers.
Future work

- Explore the time course of dialectal information integration
- Look at the role of prescriptive norms on response strategies
- Test the responses of speakers in a merged region undergoing potential demerger
  - Participants will have had mixed input
  - Here, [s] and [θ] both form part of the local sociolinguistic repertoire
- Stay tuned!
Thank you! ¡[gracias]!

Meredith Tamminga
Talkers and participants
Audience at Common Ground at Penn


Table 1: Mixed model

<table>
<thead>
<tr>
<th>Dependent variable: Reaction time in seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes: (/θ/)</td>
</tr>
<tr>
<td>Yes: (/θ/); Mexican</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Yes: (/s/); Mexican</td>
</tr>
<tr>
<td>(θ/)</td>
</tr>
<tr>
<td>(θ/); Mexican</td>
</tr>
<tr>
<td>CVCV</td>
</tr>
<tr>
<td>Log frequency</td>
</tr>
<tr>
<td>Block 2</td>
</tr>
<tr>
<td>(/s/)</td>
</tr>
<tr>
<td>Ultimate stress</td>
</tr>
<tr>
<td>Yes: Mexican</td>
</tr>
<tr>
<td>Yes: (/s/); Mexican</td>
</tr>
<tr>
<td>Yes: (/s/)</td>
</tr>
<tr>
<td>CVCVC</td>
</tr>
<tr>
<td>Mexican</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>Log Likelihood</td>
</tr>
<tr>
<td>Akaike Inf. Crit.</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2: Mixed linear effects model made in R with the Lme4 package.
## Test stimuli

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Log frequency</th>
<th>Gloss</th>
<th>Stimulus</th>
<th>Log frequency</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>yeso</td>
<td>2.274157849</td>
<td>cast</td>
<td>yace</td>
<td>2.330413773</td>
<td>lay</td>
</tr>
<tr>
<td>sigo</td>
<td>3.468199586</td>
<td>I follow</td>
<td>zona</td>
<td>3.457427693</td>
<td>zone</td>
</tr>
<tr>
<td>silla</td>
<td>3.363235804</td>
<td>chair</td>
<td>cine</td>
<td>3.388988785</td>
<td>cinema</td>
</tr>
<tr>
<td>liso</td>
<td>1.681241237</td>
<td>smooth</td>
<td>roce</td>
<td>1.785329835</td>
<td>scrape</td>
</tr>
<tr>
<td>suma</td>
<td>2.707570176</td>
<td>sum</td>
<td>cima</td>
<td>2.850033258</td>
<td>peak</td>
</tr>
<tr>
<td>sana</td>
<td>2.589949601</td>
<td>healthy</td>
<td>zorro</td>
<td>2.539076099</td>
<td>fox</td>
</tr>
<tr>
<td>vaso</td>
<td>3.035429738</td>
<td>cup</td>
<td>doce</td>
<td>3.085647288</td>
<td>twelve</td>
</tr>
<tr>
<td>subo</td>
<td>2.484299839</td>
<td>I climb</td>
<td>cera</td>
<td>2.434568904</td>
<td>wax</td>
</tr>
<tr>
<td>sota</td>
<td>1.342422681</td>
<td>joker</td>
<td>cepo</td>
<td>1.230448921</td>
<td>stocks</td>
</tr>
<tr>
<td>sabor</td>
<td>2.910090546</td>
<td>flavor</td>
<td>civil</td>
<td>2.930439595</td>
<td>civil</td>
</tr>
<tr>
<td>grasa</td>
<td>2.796574333</td>
<td>fat (n)</td>
<td>plaza</td>
<td>2.802089258</td>
<td>plaza</td>
</tr>
<tr>
<td>cosa</td>
<td>4.404508729</td>
<td>thing</td>
<td>dice</td>
<td>4.37697822</td>
<td>say</td>
</tr>
<tr>
<td>brisa</td>
<td>2.334453751</td>
<td>breeze</td>
<td>lazo</td>
<td>2.396199347</td>
<td>knot</td>
</tr>
<tr>
<td>pesa</td>
<td>2.586587305</td>
<td>weigh</td>
<td>trece</td>
<td>2.617000341</td>
<td>thirteen</td>
</tr>
<tr>
<td>pisa</td>
<td>2.187520721</td>
<td>step (v)</td>
<td>tiza</td>
<td>2.158362492</td>
<td>chalk</td>
</tr>
<tr>
<td>prisa</td>
<td>3.33304403</td>
<td>hurriedness</td>
<td>brazo</td>
<td>3.365300749</td>
<td>arm</td>
</tr>
</tbody>
</table>