Gradient effects of L1 dialect on L2 lexical processing

Christine Shea
Department of Spanish and Portuguese
University of Iowa
There has been considerable work on how L1 speakers process and adapt to different dialects of their native language.
Dialect Processing

- Short-term: need to deal with the variability in the speech stream that arises from exposure to a new dialect

- Long-term: adaption to the new variety occurs

- Also important to keep in mind the difference between processing and producing different dialects

- And the prestige enjoyed by each dialect in the broader community
L1 Dialect Adaptation

- Sumner and Samuel (2009): examined speakers, overhearers and non-speakers of r-less English dialects
  - speakers of the non-dominant dialect can process words in both dialects faster than those of the dominant dialect

- Non-dominant dialect speakers: more flexible in lexical recognition tasks than those of the dominant dialect

- not dependent upon use
• There is also extensive work on how second language learners acquire certain dialect features after prolonged exposure to the L2 (usually study abroad)
• This raises issues of what a base dialect is in L2 acquisition
We turn this issue around and ask how L1 dialect differences may affect the acquisition of an L2
L1 Dialect affects L2 Acquisition

- extensive work by Escudero and her collaborators on vowel acquisition

- shows that L1 and L2 target dialects play a role in how vowel categories are assimilated cross-linguistically
Current Study

- Set the issue up in terms of what exactly is transferred in this type of situation

- Linguistics traditionally considers speakers of different dialects as having different grammars (constraint rankings)

- The question that then arises is in situations where listeners are perceptually ‘bidialectal’ (as in Sumner and Samuel’s study) is which dialect features are transferred to the L2 processing?
L1 Dialect affects L2 Acquisition

- Basic question informing all L2/Ln acquisition is how best to characterize the constraints on non-first language acquisition in adulthood

- Essentially revolves around the notion of transfer
  - interference from the native language
    - negative transfer – impedes learning
Current Study

- Examine how L1 dialect affects lexical processing in an L2

- I am going to contextualize the issue in terms of the notion of transfer

- ... and so what?

- Not only in terms of what the learner brings directly to the L2 (processing, structure) but also what the L2 learner expects from the L2
  - i.e., what is inferred
Current Study

- We show that L2 learners transfer L1 dialect features to L2 lexical processing tasks
  
  (what is transferred)

- These effects diminish with increased proficiency in the target language
  
  (increased experience)
Current Study: Dialect regions

Alternation: [-s] aspiration in Spanish

Dialects:

1. Colombian highland: Bogota
2. Colombian lowland (Pacific Coast, Barranquilla)
Aspiration in Spanish

One of the most extensively studied sociolinguistic variables of Spanish

[-s] aspiration occurs in non-highland Latin American Spanish and Andalusian varieties of Spain

For aspirating dialects, almost all aspire in [sC] contexts, word-internally:

- vasco ‘Basque’, masc → [bah.ko]
- vaso ‘glass’ → [ba.so]
- pasto ‘grass’ → [pah.to]
- pato ‘duck’ → [pa.to]
Method: auditory form priming with lexical decision
Auditory Form Priming with Lexical Decision

**Prime**

“historia”
i[h]toria

**Target**

ISI: 500 ms

“historia”
i[h]toria

**REAL WORD**

**NON-WORD**

Lexical decision tasks: words and nonwords are presented, participants judge if target is an existing word

**Dependent Variable:** Reaction Time to target
Auditory Form Priming

Pre-exposure to related stimuli speeds up recognition/retrieval of target stimuli

\[ \text{[Target]} = \text{[Prime]} \]

MAXIMUM PRIMING EFFECT

Priming effect reflects the preactivation of the target word, based upon the degree of perceived similarity
Auditory Form Priming

- Allows us to probe the stored phonetic form

**Thus:**

If [-h-] can prime [-s-], then the listener recognizes them as ‘equal’, or at least equally good exemplars of the target word.
## Spanish Trials

<table>
<thead>
<tr>
<th>Condition</th>
<th>Prime</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IDENTITY</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SAME ASPIRATION</strong></td>
<td>fui[h]te</td>
<td>fui[h]te</td>
</tr>
<tr>
<td><strong>DIFFERENT ASPIRATION</strong></td>
<td>ha[s]ta</td>
<td>ha[h]ta</td>
</tr>
<tr>
<td>Condition</td>
<td>Prime</td>
<td>Target</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>IDENTITY</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SAME ASPIRATION</strong></td>
<td>fui[h]te</td>
<td>fui[h]te</td>
</tr>
<tr>
<td><strong>DIFFERENT ASPIRATION</strong></td>
<td>ha[s]ta</td>
<td>ha[h]ta</td>
</tr>
<tr>
<td><strong>CONTROL</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SAME ASPIRATION</strong></td>
<td>a[h]co</td>
<td>a[h]co</td>
</tr>
<tr>
<td><strong>DIFFERENT ASPIRATION</strong></td>
<td>ba[s]ton</td>
<td>fra[h]co</td>
</tr>
</tbody>
</table>
Spanish stimuli: recorded by two female speakers of Bogota Spanish and Cartagena/Barranquilla Spanish

Aspirated segments were between 20-50ms
- this reflects the measurements of the speakers themselves
- samples were naturalistic
Lexical items were selected from the NIM online data base
Procedure

- 30 participants for each dialect group
- Tested in their hometowns
- Recruited from undergraduate populations at private (Bogota) and public universities (Barranquilla)
RESULTS:
Spanish Form Priming
Spanish IDENTITY TRIALS
Prime-Target: significant ([F(3, 61)=6.170, p<0.001

Group: p=0.003; Pairwise Comparisons (Bonferroni Adjusted, p<0.05)
DIFFERENCE SCORES BY GROUP - SPANISH

CONTROL (1) - IDENTITY: consistent aspiration

[[-h]:[-h]]

BARR

BOG

[[-s-]:[-s]]

BARR

BOG

CONTROL (2) - IDENTITY: inconsistent aspiration

[[-h]:[s]]

BARR

BOG

['calle'- 'fresa'

vs.

'mae[h]/[s]tro'- 'mae[s]/[h]tro']

[[-s-]:[-s]]

BARR

BOG
Discussion: Spanish

(1) Barranquilla listeners: demonstrate greater flexibility when listening to a different dialect of their L1

(2) Bogota listeners: less flexible
Discussion: Spanish

- L1 dialect was significant
- And there was an interaction between L1 dialect and trial-type
Experiment 2: English priming

- Same procedure
- Same participants
- English stimuli: recorded by two native female speakers of standard North American English

Stimuli:

Half were [-s], half [-h]
Critical Predictions

(1) There will be a main effect for L1 dialect
(1) Barranquilla listeners: will demonstrate equal flexibility when listening to the modified English forms
Stimuli: English form priming

Aspiration was between 20-40ms
RESULTS:
English Form Priming
ENGLISH IDENTITY TRIALS

Prime Condition: significant ([F(3, 61)=4.22, p<0.001]

Avg Reaction Time

Pairwise Comparisons (Bonferroni Adjusted, p<0.05); Group-n.s.
DIFFERENCE SCORES BY GROUP (English)

CONTROL (1) – IDENTITY: consistent aspiration

CONTROL (2) – IDENTITY: inconsistent aspiration
L1 Dialect\textsubscript{1}: Bogota

\([-s]\sigma\) coda

L1 Dialect\textsubscript{2}: Barranquilla

\([-s]\sigma\) coda/
\([-h]\sigma\) coda

ho[h]pital

L2
Summary

- Barranquilla listeners are faster at recognizing English trials with different prime-target aspiration patterns.

- Now: what about off-line knowledge?

- We propose that experience with English will serve to modulate the expectations of the Barranquilla listeners and affect what it is they are transferring to their target language.
Mean Log10 RT for Different Aspiration

Vocabulary Z-Score

$r = -0.56$

$t = -6.23$, $p < 0.05$

$0.95$ C.I. = $-0.87$, $-0.55$
Aspirating Dialect  Non-Aspirating Dialect

Barranquilla

Non-Aspirating Dialect

Bogota

Aspirating Dialect  Non-Aspirating Dialect

PRIME [asta]  TARGET [ahta]

Consistent

Inconsistent

38
Aspirating Dialect

Non-Aspirating Dialect

Barranquilla

Non-Aspirating Dialect

Aspirating Dialect

Bogota

Consistent

Inconsistent
Another possibility: The Barranquilla speakers are using their relatively greater perceptual flexibility to adapt to the modified English tokens as well.

- For example, if we modified the English tokens in any way, not only in the way that corresponded to their L1 dialect expectations, they would process it faster than the Bogotá listeners.
- We would argue that no, this does not hold.
- The non-word controls could possibly provide evidence against this.
- The Barranquilla listeners were no faster or accurate on those tokens than the Bogotá listeners.
Integration vs. Prediction

This also addresses the issue of whether learners are actually actively expecting the aspirated form or merely tolerating both forms, i.e., the prediction/integration debate.

The data presented here suggests that prediction, or expectations are at play because there is no reason why the English aspirated forms should be easier to process if the aspiration expectation were not active.

They have never heard this, they should not find it easier to integrate unless it follows listener expectations.
Conclusions

- Dialects, languages, etc., cluster in terms of their similarity, and learning a new language involves inferring its relationship with previously established clusters.

- Strong prior beliefs about how things are clustered can hinder learning.

- New linguistic structures will be learned only if it can overcome L1 clusters.

- This important because it does not assume that everything from the L1 will necessarily be transferred.

- Elements eligible for transfer are those that are most similar to between the L1 and the target language.

- Various things can affect what is transferred – spelling, explicit instruction in the case of L2 learners.
Conclusions

- In closing, prior knowledge and prior expectations guide how new linguistic input will be processed, especially at the beginning, when evidence from the target language is sparse.

- Influence of prior knowledge diminishes as proficiency with the target language increases (off-line RT measurements).

- But it might never completely disappear.
¡Gracias!

Thanks to the University of Iowa Old Gold Fellowship for funding the research trip to Colombia.

Thanks also to students at the Universidad Javeriana (Bogota) and the Universidad del Atlántico (Barranquilla)
Transfer

- idea of transfer

- work through notion of variability,

- here: dialect variability, generalized and applied to the L2

- Shows effect of experience and shows ability to generalize

- but not permanent – reflects experience with the L2, USE the forms in their own speech AND also have stayed in the speech community the longest, racial identification with the dialect as well

- do not produce it, but dialect-consistent
Given the time constraints under which comprehension operates, one clear benefit of being able to predict upcoming material is that it may allow a listener or reader to produce an overt response more quickly, without waiting for the material itself to become available.
More recent investigations of anticipatory processing have been spurred, in part, by interest in questions of how generalized prediction might be, what it might be used for, what information is available to predictive processing, and what aspects of upcoming input are being predicted.
We will now present evidence for the “strong” form of prediction—that is, experimental evidence showing that specific linguistic items or their features are pre-activated at some time point prior to encountering the confirmatory bottom-up input. Such evidence turns out to be somewhat tricky to obtain, due to the challenge of identifying processing related to an item that has not yet occurred. In particular, one difficulty centers on the prediction/integration debate—in other words, the challenge of distinguishing facilitation effects at a target word as being due to that word being predicted versus it being simply easier to integrate upon receipt.
## Log10 Latencies English (Only for Same Trials)

**Reference:** Barranquilla, ‘h’, ‘Identity’

### Fixed Effects:

<table>
<thead>
<tr>
<th>Effect</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t Value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>3.38</td>
<td>0.019</td>
<td>179</td>
<td>0.0002</td>
</tr>
<tr>
<td>Group (Bogota)</td>
<td>-0.05</td>
<td>0.02</td>
<td>2.13</td>
<td>0.029</td>
</tr>
<tr>
<td>Unmatched</td>
<td>0.06</td>
<td>0.01</td>
<td>4.64</td>
<td>0.00001</td>
</tr>
<tr>
<td>Bogota * ‘s’</td>
<td>-.063</td>
<td>0.024</td>
<td>-2.62</td>
<td>0.0088</td>
</tr>
</tbody>
</table>
Linear Mixed Model in R (lme4, Bates et al., 2014)

- Used R (R Core Team, 2012) and lme4 (Bates, Maechler & Bolker, 2012) to perform a linear mixed effects analysis of the relationship between Reaction Time and group, aspiration and trial type

Fixed effects: Reaction Time, Group, Aspiration, Trial Type

Random effects: Participant, Trial
**RT Spanish: Identity Trials**  
*(Log10 RT values)*  
**Reference:** Barranquilla, ‘h’, ‘Identity’ ‘Matched’

**Fixed effects:**

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>T Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>3.36</td>
<td>.02</td>
<td>127.5</td>
<td>0.0001</td>
</tr>
<tr>
<td>GROUP Bogota</td>
<td>0.05</td>
<td>0.012</td>
<td>3.95</td>
<td>0.001</td>
</tr>
<tr>
<td>[-s]</td>
<td>-0.07</td>
<td>0.04</td>
<td>-2.00</td>
<td>0.0465</td>
</tr>
<tr>
<td>UNMATCHED [-h]-[s]</td>
<td>0.08</td>
<td>0.011</td>
<td>5.27</td>
<td>0.0001</td>
</tr>
<tr>
<td>Bogota* [-s]</td>
<td>-0.043</td>
<td>0.015</td>
<td>-2.77</td>
<td>0.03</td>
</tr>
<tr>
<td>Bogota * [h]-[s]</td>
<td>0.023</td>
<td>0.015</td>
<td>-2.01</td>
<td>.0402</td>
</tr>
</tbody>
</table>
We tested the full model vs. the null model (with only Group as fixed effect)

Significant difference in likelihoods ($\chi^2(14)=137.02$, $p=0.00022$),
Model Significance

- We tested the full model vs. the null model (with only Group as fixed effect)

Significant difference in likelihoods ($\chi^2(14)=137.02 \ p=0.0000913$)
In general, the Barranquilla listeners are better at processing the [-h] variant than the Bogota listeners.

Holds for both English and Spanish.

Perceived structural similarities between English and Spanish are inferred.
These issues speak directly to the nature of representations in L1:

If both dialects are stored, representations are necessarily detailed and flexible – reflect experience with a dialect
  • supports more exemplar-based approaches

L2:

If dialect-specific differences are noted in L2 acquisition, some sort of abstract knowledge might also be involved in lexical processing
Sample Questions from Dialect Questionnaire

I want to sound like speakers from ____________.

I do NOT want to sound like speakers from ________________.

The ‘best’ Spanish is spoken by people from ________________.

The ‘poorest’ Spanish is spoken by people from ________________.

The easiest Spanish accent to understand is ________________.

The most difficult Spanish accent to understand is ________________.

While it is obvious I am not a native Spanish speaker, when I speak Spanish I have an accent from ________________.

When we say someone speaks with the ‘ceseo’ or ‘seseo’ or ‘distinción’ it means:
### LOG10 LATENCIES ENGLISH (ONLY FOR SAME TRIALS)

**Reference:** Barranquilla, ‘h’, ‘Identity’

#### Fixed Effects:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>3.38</td>
<td>0.019</td>
<td>179</td>
<td>0.0002</td>
</tr>
<tr>
<td>Group (Bogota)</td>
<td>-0.05</td>
<td>0.02</td>
<td>2.13</td>
<td>0.029</td>
</tr>
<tr>
<td>Unmatched</td>
<td>0.06</td>
<td>0.01</td>
<td>4.64</td>
<td>0.00001</td>
</tr>
<tr>
<td>Bogota * ‘s’</td>
<td>-0.063</td>
<td>0.024</td>
<td>-2.62</td>
<td>0.0088</td>
</tr>
</tbody>
</table>
Used R (R Core Team, 2012) and lme4 (Bates, Maechler & Bolker, 2012) to perform a linear mixed effects analysis of the relationship between Reaction Time and group, aspiration and trial type.

Fixed effects: Reaction Time, Group, Aspiration, Trial Type

Random effects: Participant, Trial
### RT Spanish: Identity Trials

(Log10 RT values)


<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>T Value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>3.36</td>
<td>.02</td>
<td>127.5</td>
<td>0.0001</td>
</tr>
<tr>
<td>GROUP Bogota</td>
<td>0.05</td>
<td>0.012</td>
<td>3.95</td>
<td>0.001</td>
</tr>
<tr>
<td>[-s]</td>
<td>-0.07</td>
<td>0.04</td>
<td>-2.00</td>
<td>0.0465</td>
</tr>
<tr>
<td>UNMATCHED [-h]-[s]</td>
<td>0.08</td>
<td>0.011</td>
<td>5.27</td>
<td>0.0001</td>
</tr>
<tr>
<td>Bogota* [-s]</td>
<td>-0.043</td>
<td>0.015</td>
<td>-2.77</td>
<td>0.03</td>
</tr>
<tr>
<td>Bogota * [h]-[s]</td>
<td>0.023</td>
<td>0.015</td>
<td>-2.01</td>
<td>.0402</td>
</tr>
</tbody>
</table>
Model Significance

- We tested the full model vs. the null model (with only Group as fixed effect)

Significant difference in likelihoods ($\chi^2(14)=137.02$ $p=0.00022$),
Model Significance

- We tested the full model vs. the null model (with only Group as fixed effect)

Significant difference in likelihoods ($\chi^2(14)=137.02 \ p=0.0000913$)
In general, the Barranquilla listeners are better at processing the [-h] variant than the Bogota listeners.

Holds for both English and Spanish.

Perceived structural similarities between English and Spanish are inferred.
DIFFERENCE SCORES BY GROUP - SPANISH

CONTROL (1)-FULL IDENTITY: consistent aspiration

[-h]:[-h]

'BARR BOG'

'n:300
'500'

CONTROL (2) - IDENTITY : inconsistent aspiration

[-h]:[s]

'BARR BOG'

'n:300
'500'

'casa'- 'mesa'
vs.
'pa[h][s]to'- 'pa[h/[s]]to'

'calle'- 'fresa'
vs.
'mae[h]/[s]tro'- 'mae[s]/[h]tro'