Substantive bias and the acquisition of final (de)voicing patterns

Eleanor Glewwe  Jesse Zymet
Jacob Adams  Rachel Jacobson  Anthony Yates  Ann Zeng
Robert Daland
University of California, Los Angeles

1. Introduction

A typological asymmetry:

- Distribution of voicing in obstruents (e.g., Yu 2004):
  - All languages allow voiceless obstruents.
  - Voiced obstruents: none, contextually predictable, contrastive.
- Word-finally
  - A voicing contrast is often neutralized (Dutch, Russian, Catalan, Turkish, etc.)
  - Neutralization is always to voiceless obstruents, with one exception (monosyllabic nouns in Lezgian; Yu 2004).
- Is this because of a **substantive learning bias**?
  - Proposes that learners are biased to learn phonetically natural patterns (Wilson 2006 et seq);
  - Final devoicing facilitates articulatory ease (e.g., Westbury & Keating 1986), and so we would expect learners to favor it over final voicing.
- Alternative: **channel bias**.
  - Proposes that physics of speech production/soundwave transmission, psychoacoustics of speech perception make final devoicing more likely to arise diachronically than final voicing (Hyman 1976; Ohala 1992, 1993).

Moreton & Pater (2012a,b) single out two learning biases proposed in prior literature:

- **Substantive bias** (aka naturalness)
  - Easier to learn patterns that are phonetically motivated.
- **Complexity bias** (aka formal complexity)
  - Easier to learn patterns that require fewer features to state.
- **Their findings**: Complexity bias is widely confirmed by the experimental literature, but the evidence for substantive bias is weak or inconclusive.

We test for substantive bias and complexity bias in an artificial language learning experiment. Participants are exposed to one of three artificial patterns:

- **Final devoicing** pattern;
- **Final voicing** pattern;
- A more complex **exchange rule** pattern (Anderson & Browne 1973, Moreton 2004), which involves both final devoicing and final voicing.
Substantive bias and the acquisition of final (de)voicing
LSA 2018

Predictions:
• Substantive bias: Devoicing should be learned better than Voicing.
• Complexity bias: Exchange should be learned worse than both Devoicing and Voicing.

Our results, in short:
• Voicing is learned better than Devoicing.
• Exchange is not learned worse than Devoicing and Voicing.

Implications:
• We propose to narrow the substantively biased learning hypothesis.
• Finley & Badecker (2008) (emphasis ours): hypothesis predicts that learners are biased to form grammars that “maximize articulatory ease and perceptual salience”.
• Our results suggest that articulatory ease does not bias learning. Other recent experiments support our view.
• Instead, processes rooted in articulatory ease arise through channel bias.
• Suggests markedness constraints motivated by articulatory ease are induced, not innate.

2. Experiment

• Artificial grammar learning experiment
  o Familiarized to plural-singular alternation
  o Test: hear plural, select singular (two-alternative forced choice task), half familiar, half novel

• Conditions
  o Final devoicing
  o Final voicing
  o Exchange (complex pattern involving final devoicing and final voicing)

2.1 Methods
2.1.1 Participants

• Native English speakers with no/negligible proficiency in any language with active word-final devoicing
• Recruited from the UCLA Psychology Subject Pool
• 22 for Devoicing, 23 for Voicing, and 20 for Exchange
Substantive bias and the acquisition of final (de)voicing
LSA 2018

2.1.2 Stimuli

- Stimuli categorized by condition:

<table>
<thead>
<tr>
<th>Condition</th>
<th>Baseline</th>
<th>Non-alternating</th>
<th>Alternating</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEVOICING</td>
<td>UR sonorant</td>
<td>UR voiceless stop</td>
<td>UR voiced stop</td>
</tr>
<tr>
<td></td>
<td><em>komál-</em></td>
<td><em>tuláp-</em></td>
<td><em>mulé-</em> (DEVOICING)</td>
</tr>
<tr>
<td>VOICING</td>
<td>UR sonorant</td>
<td>UR voiced stop</td>
<td>UR voiceless stop</td>
</tr>
<tr>
<td></td>
<td><em>komál-</em></td>
<td><em>mulé-</em></td>
<td><em>tuláp-</em> (VOICING)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>Baseline</th>
<th>Alternating</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXCHANGE</td>
<td>UR sonorant</td>
<td>UR voiced stop (DEVOICING)</td>
</tr>
<tr>
<td></td>
<td><em>komál-</em></td>
<td><em>mulé-</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>tuláp-</em> (VOICING)</td>
</tr>
</tbody>
</table>

- Auditory stimuli
  - CVCVC stems with plural suffix [-i]
  - Vowels drawn from [i e a o o u], sonorants from [m n l ɹ], obstruents from [p t k b d g]
  - Recorded by a phonetically trained female native speaker of American English.

- Visual stimuli
  - Multiplied and resized to make plural picture.

2.1.3 Design

- Familiarization
  - See plural picture & hear plural form;
  - Then see singular picture & hear singular form.
  - 12 sonorant-final stems, 12 voiceless-final stems, 12 voiced-final stems.

- Test
  - See plural picture & hear plural form;
  - Then see singular picture & hear two singular forms, one alternating and one non-alternating:
    - For obstruent-final stems, one choice voiced and one choice voiceless;
    - For sonorant-final stems, *r ↔ l* and *m ↔ n*.
  - Press '1' for first and '2' for second.
  - Tested on 12 (6 familiar, 6 novel) of each stem type (sonorant, voiceless, voiced).
Substantive bias and the acquisition of final (de)voicing
LSA 2018

TRAINING PHASE (voicing condition)

Repeat the word you heard.
Then, press any key to continue.

TESTING PHASE (voicing condition)

Press any key to continue.

Which word is the correct word for this picture?
Press “1” for the first word.
Press “2” for the second word.
Substantive bias and the acquisition of final (de)voicing
LSA 2018

2.2 Results

Figure 1 shows the rates of correct response by Condition and Stem Type, with Familiar stems (heard in training) and Novel stems (not heard in training) pooled.

• Stem Type refers to final underlying consonant of the stem.
• A indicates stem types that are alternating in a given condition.

First impressions:
• Across all three conditions, non-alternating sonorant-final stems learned well (correct response rate > 50%).
• In Devoicing, non-alternating voiceless-final stems learned well, but alternating, voiced-final stems not learned (correct response rate not different from chance, i.e. 50%).
• In Voicing, non-alternating voiced-final stems learned well, and alternating, voiceless-final stems learned fairly well (correct response rate significantly above chance).
• In Exchange, alternating (voiceless-final and voiced-final) stems learned also (correct response rate significantly above chance).

Analyses:
• Results analyzed with mixed-effects logistic regressions
• Dependent variable always Response (correct/consistent vs. incorrect/inconsistent)
• All models included random intercepts for Subject and Item
Substantive bias and the acquisition of final (de)voicing
LSA 2018

Full model with predictors **Condition** (D, V, or E) and **Alternating** (Yes or No) (see Appendix for regression model):

- Reference levels: Devoicing and Non-Alternating
- Main effect of Condition such that **Voicing > Devoicing**: overall performance in Voicing condition is significantly better than in Devoicing condition ($p = 0.01$)
- Main effect of Alternating such that **Alternating < Non-Alternating**: Participants performed worse on alternating stems ($p < 0.001$)
  - But a significant interaction whereby performance is not as bad on Alternating stems in Exchange ($p < 0.001$)

Within-condition models with predictor **Alternating** (Yes or No):

- In Devoicing, **Alternating < Non-Alternating** ($p < 0.001$)
- In Voicing, **Alternating < Non-Alternating** ($p < 0.001$)
- In Exchange, no significant main effect of Alternating (Alternating < Non-Alternating; $p = 0.22$)

### 3. Discussion

Interim summary:

- Final devoicing is learned **no better** than final voicing (in fact, the latter is learned better);
  - This result supports one of Moreton & Pater (2012a,b)’s proposals: no evidence for substantive bias.
- The exchange pattern is **not** learned worse than Devoicing and Voicing; it is learned differently.

Why was unnatural Voicing learned better than natural Devoicing?

- Hard to say, but [tuláb], with stressed final syllable with a final voiced obstruent, may be phonotactically rare in English, thus [tuláp-i] ~ [tuláb] alternation particularly noticeable.
- Albright & Do (2017)’s experiment: English-speaking participants reacted in particular to phonotactically illicit sequences with [ŋ], which they attribute to salience.
- Alternatively, there may be a complexity bias story:
  - Voicing participants may have grouped sonorants and voiced obstruents together as [+voice] and learned the constraint *[voice] to drive their alternation.
  - If sonorants are [+voice], Devoicing participants would have had to learn the more complex constraint *[-son, +voice] to drive their alternation.
  - If this is the case, Voicing > Devoicing supports complexity bias.

Why was the complex exchange pattern learned so well?

- Exchange participants may have performed well on Alternating stem types because they were exposed to twice as many alternating stems as Voicing and Devoicing participants.
- Exchange participants may have simply learned “alternate 2 out of 3 items”.

How should these results be interpreted?
3.1 Narrowing substantively biased learning theory

We single out an illustrative quote from Finley & Badecker (2008) (emphasis ours):

“The substantively biased theory of learning … hypothesizes that learning biases shape the distribution of linguistic patterns across the world’s languages. The easiest patterns to learn are the ones most common cross-linguistically. Patterns that are phonetically grounded… are [amongst] the easiest to learn, [and thus] are most likely to appear cross-linguistically. For example, learners are biased to form grammars that

**maximize articulatory ease and perceptual salience.”**

Given the above statement and prior findings that pattern complexity plays a role in learning (Moreton & Pater 2012a,b), we get the following graph:

![Current State of Biased Learning Theory](image)

But final devoicing facilitates articulatory ease (Jaeger 1978, Ohala 1983, Westbury & Keating 1986), and our findings suggest final devoicing is no more readily learned than final voicing.

**Our proposal:** Articulatory ease does not bias learning.

![Modified Biased Learning Theory](image)

- Learner searches space of natural classes of sounds.
- She may identify featurally simple patterns with less difficulty (complexity bias),
- and perceptual factors may lead her to identify some patterns with less difficulty than others (“substantive bias” — that is, *perceptual bias*),
- but no sense that she identifies patterns maximizing articulatory ease with less difficulty.
Substantive bias and the acquisition of final (de)voicing
LSA 2018

Converging evidence toward this proposal:

Skoruppa & Peperkamp (2011): artificial vowel disharmony is learned as readily as vowel harmony.
- Despite vowel harmony being relatively common,
- and despite claims of articulatory benefit of vowel harmony (Benus 2005).

Do, Zsiga & Havenhill (2016): artificial post-nasal devoicing learned as readily as post-nasal voicing.
- Despite post-nasal voicing being relatively common,
- and despite claims of articulatory benefit of post-nasal voicing (Hayes & Stivers 1995).

Our experiment: final voicing is learned more readily than final devoicing.
- Despite final devoicing being relatively common,
- and despite claims of articulatory benefit of final devoicing (Westbury & Keating 1986).

Then why are patterns that maximize articulatory ease so common?
- **Channel bias**: Subtle phonetic patterns that facilitate ease of articulation develop over time and are then phonologized (Moreton & Pater 2012a,b).

This suggests strict separation between synchrony and diachrony:
- Phonetically grounded patterns develop over time and phonologize, which explains some differences in typological frequency;
- But this does not mean that learners are biased to form synchronic grammar patterns that maximize articulatory ease.

What does this mean for constraint-based theory? Are constraints induced or are they innate?
- If our proposal is correct, then this would suggest that lots of constraints—e.g., markedness constraints motivated by articulatory ease—are induced (Hayes & Wilson 2008; Alderete, Tupper & Frisch 2013).
- Experiments show that phonotactic constraints can be rapidly induced (Onishi et al. 2002, Chambers et al. 2003, Warker et al. 2008, Linzen & Gallagher 2017).
- But maybe not everything about constraints is induced:
  - Insofar as perceptual factors bias learning, we might expect some constraints to be present in CON in the initial state, and for some rankings to be more or less fixed:
  - Constraints against drastically unfaithful mappings;
  - Fixed rankings of faithfulness constraints set by the P-Map (Zuraw 2013, Hayes & White 2015; cf. Steriade 2001);

3.2 Can we just discard substantively biased learning altogether?

Other experimental results also cast doubt on substantive bias (cf. Moreton & Pater 2012a,b):
- Greenwood (2016): A language with only voiceless obstruents word-finally is not learned better than a language with only voiced obstruents word-finally.
Substantive bias and the acquisition of final (de)voicing
LSA 2018

• Glewwe (in press): When a learner is exposed to and acquires a contrast in coda position, they do not necessarily infer a contrast in onset position.

**Problem:** Some experimental results suggest that perceptual factors *do* play a role in learning.

Wilson (2006): Learners exposed only to k → tʃ/ _e apply rule before *e* and *i*, but learners exposed only to k → tʃ/ _i apply rule *only before i*.

- Mirrors typological generalization that palatalization before [e] asymmetrically implies palatalization before [i].
- Exposed to a perceptually distant mapping, learners infer analogous perceptually less distant mappings, but not vice versa

White (2014), with similar findings: learners are biased against acquiring saltatory patterns, in which one sound alternates with a phonetically distant sound, but intermediates do not.

- Mirrors typology: there are few attested saltatory patterns across languages, and those that are known arose by a set of historical changes (Hayes & White 2015).
- Again suggests perceptual factors play a role in learning.

Finley (2012), Kimper (2016): Participants more readily acquire artificial rounding harmony patterns with mid vowels than with high vowels.

- Mirrors generalization that mid vowels are typical triggers of rounding harmony.
- Prior explanation: roundness harder to perceive in mid vowels relative to high vowels; greater perceptual reward for harmonizing from a mid vowel (Terbeek 1977).

Overall, might be best not to throw the baby out with the bathwater; more work should narrow in on what aspects of perception bias learning. We can break up substantive factors further:

![Tree Diagram](#)

**Figure 4:** further modified biased learning theory
4. Conclusion

Artificial language learning experiment tested for whether phonetically natural final devoicing is learned better than unnatural final voicing and the unnatural and complex exchange pattern.

- Voicing was learned better than Devoicing; Exchange was learned surprisingly well.

Results support proposals put forth in Moreton & Pater (2012a,b), in that we find no evidence for substantive bias.

Experimental results conflict regarding substantively biased learning, but our experiment along with a few others suggests a divide between perceptual grounding and articulatory grounding.

- In particular, learners are not biased to form grammars that facilitate articulatory ease.

Appendix

Mixed-effects logistic regression of experiment results carried out in R with lme4 package (Bates et al. 2011)

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.771</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>Condition = Exchange (vs. Devoicing)</td>
<td>-0.242</td>
<td>0.513</td>
</tr>
<tr>
<td>Condition = Voicing (vs. Devoicing)</td>
<td>0.875</td>
<td>0.014*</td>
</tr>
<tr>
<td>Alternating = Yes (vs. No)</td>
<td>-2.061</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>Condition = Exchange × Alternating = Yes</td>
<td>1.623</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>Condition = Voicing × Alternating = Yes</td>
<td>0.104</td>
<td>0.741</td>
</tr>
</tbody>
</table>

Acknowledgments

We are grateful to Yu Tanaka, who was an early collaborator on the project, as well as to our helpful research assistants, Josh Sciortino and Rick DeAmicis. Finally, we are grateful to members of the UCLA Phonology Seminar for invaluable feedback.

References


Substantive bias and the acquisition of final (de)voicing
LSA 2018


Substantive bias and the acquisition of final (de)voicing
LSA 2018


