On the learnability of saltatory phonological alternations

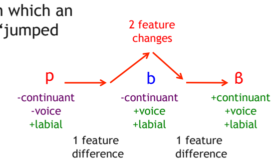
James White
Department of Linguistics, UCLA
jameswhite@ucla.edu

1. Background and research questions

- How do people learn the phonological alternations of their language?
  - Is phonological learning the result of an unbiased search for patterns?[1]
  - Or is learning guided by biases against certain patterns?
    - What types of patterns are dispreferred by learners?
    - What is the nature of these biases?
    - How can we account for them in learning models?
- Much of the recent literature supports a soft bias approach: certain patterns are dispreferred, but may be learned given enough input.[2-5]

2. Case study: Saltatory alternations

- Saltatory alternation: An alternation in which an intermediate non-alternating sound is “jumped over.”
- E.g. Campidanian Sardinian:[6]
  - /pani/ → [sːu βãi] ‘the bread’
  - /binu/ → [sːu βĩu] ‘the wine’



- Why are they interesting?
  - Relatively uncommon.
  - It has long been argued that alternations between more similar sounds are better than alternations between less similar.[7-8]
  - Accounting for them is not theoretically straightforward (e.g., classical OT[9] cannot).

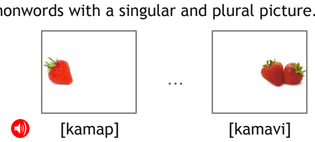
Acknowledgments

I would like to thank Bruce Hayes, Megha Sundara, Robert Daland, Kie Zuraw, Sharon Peperkamp, Marc Garellek, Karen Campbell, Adam Albright, Jennifer Cole, audiences at the UCLA phonology seminar and LSA, and anonymous LabPhon reviewers for helpful comments on this project.
I would also like to thank undergraduate research assistants who have worked on parts of this project: Kelly Ryan, Kelly Nakawatase, Ariel Quist, Emily Ho, Sarah Yiu, Ben Longwill, and Maria Manukyan.
Any remaining errors are my own.
This work was funded in part by a UCLA Summer mentorship fellowship.

3. The experiments: Are learners biased against saltatory alternations?

Method

- Artificial language learning paradigm
- Participants: 20 UCLA undergrads per experiment.
- Procedure – 3 phases:
  - Exposure: Learn alternations by hearing pairs of nonwords with a singular and plural picture.

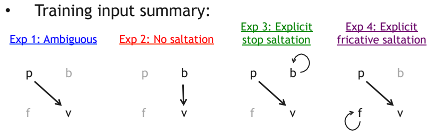


  - Verification: Tested on subset of exposure pairs. Participants choose between changing and non-changing auditory plural options (order counterbalanced).

    e.g. [kamap] ... [kamavi] -or- [kamapi]

--Repeat Exposure/Verification until 80% accuracy--

  - Generalization: Same test on novel nonwords, including some ending in untrained sounds.
- Training input summary:



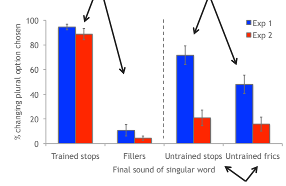
- Tested in Generalization phase on nonwords ending in [p, b, f] for each experiment.
- Also included: analogous coronals [t, d, θ, ð].
- Equal number of non-alternating fillers (mostly sonorants) in training.

  E.g. [luman] → [lumani]

Results in generalization phase

Experiments 1 & 2

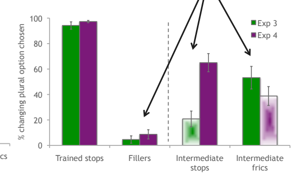
- Trained pattern: Good generalization to novel nonwords.
- Preference for changing untrained stops more than fricatives.



- Untrained sounds: Ps change intermediate sounds when ambiguously saltatory (Exp 1) much more than they do untrained non-intermediate sounds (Exp 2). = dispreference for saltatory alternations!

Experiments 3 & 4

- Explicitly saltatory cases: some learning, but a lot more errors than on fillers, despite training (compare faded bars to Fillers).



Participants learn saltatory alternations, but they are harder to learn—they continue to make errors that they don’t make for fillers.

4. Modeling overview

- How can we account for the dispreferred status, but ultimate learnability, of saltatory alternations?
- I use Maximum entropy grammar models.[10]
- Implemented using the Maxent Grammar Tool (from Bruce Hayes’s webpage).
- Provide model:
  - A set of input forms
  - For each input form, a set of candidates and the observed probability of each candidate winning
  - A set of constraints and violations
  - Prior (soft bias), 2 settings for each constraint:
    - μ = preferred weight for a constraint
    - σ^2 = how tightly a constraint weight is constrained to its μ during learning
- Training data: Same as experimental participants received.
- Model output:
  - Constraint weights
  - For each input form, the predicted probability of each candidate
- Same markedness constraints for all models:
  - *V[-voice]V = penalize intervocalic voiceless sounds
  - *V[-continuant]V = penalizes intervocalic stops

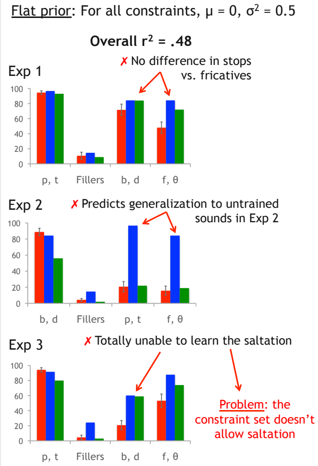
5. The models

- Exp Results
- Model Predictions
- Best Possible Fit (Model directly trained on experimental results)

I. Traditional faithfulness[9], Unbiased

Faithfulness constraints:
IDENT(voice), IDENT(cont), IDENT(son)

Flat prior: For all constraints, μ = 0, σ^2 = 0.5

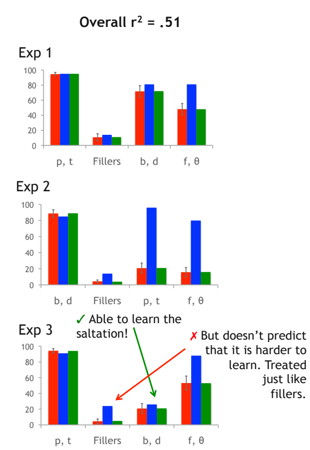


II. *MAP constraints, Unbiased

Faithfulness constraints:
More expressive *MAP constraints[11] for every pair of sounds

E.g., *MAP(p, v) = Don’t map underlying /p/ to surface [v]

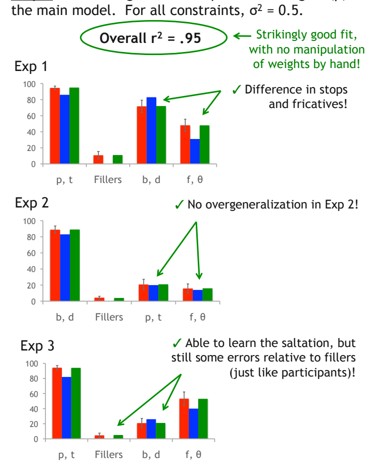
Flat prior: For all constraints, μ = 0, σ^2 = 0.5



III. *MAP constraints, Similarity bias

Faithfulness constraints: *MAP constraints

Similarity bias prior - similar to Steriade’s P-map[8]:
Step 1: Use confusion matrix probabilities[12] as input to a separate Maxent model with *MAP constraints.
Result: *MAP constraints receive weights based on confusability; Constraints penalizing more confusable (i.e., more similar) sounds get lower weights.
E.g., *MAP(b, v) has lower weight than *MAP(p, v).
Step 2: These weights become preferred weights (μ) in the main model. For all constraints, σ^2 = 0.5.



6. Conclusions

- Language learners disprefer saltatory alternations, but they are not unlearnable.
- A MaxEnt grammar model can account for both of these observations, provided that:
  - It has a more expressive set of constraints (such as the *MAP variety).
  - It has a similarity bias, derived from confusion matrix data, preferring alternations between similar sounds.

7. References

[1] Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge Univ. Press.
[2] Hayes, B., et al. (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language*.
[3] Hayes, B., & White, J. (in press). Phonological naturalness and phonotactic learning. *Linguistic Inquiry*.
[4] Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cog. Sci*.
[5] Finley, S., & Badecker, W. (2008). Substantive biases for vowel harmony languages. *WCCFL 27*.
[6] Bolognesi, R. (1998). *The phonology of Campidanian Sardinian: A unitary account of self-organizing structure*. Holland Academics.
[7] Trubetzkoy, N. S. (1939). *Principles of Phonology*.
[8] Steriade, D. (2001/2008). The phonology of perceptibility effects: the P-map and its consequences for constraint organization.
[9] Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Blackwell.
[10] Hayes, B, & Wilson, C. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*.
[11] Zuraw, K. (2007). The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog. *Language*.
[12] Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: a study of perceptual features. *JASA*.