FEATURE-BASED GENERALIZATION AND ARTIFICIAL LANGUAGE LEARNING

SARA FINLEY* & WILLIAM BADECKER, *Johns Hopkins University*

This research was supported in part by an IGERT grant from the National Science Foundation.

BACKGROUND

- Artificial grammar experiments have been used to address questions of learnability (e.g. Newport and Aslin, 2003) and naturalness (e.g. Pycha et al., 2003; Wilson, 2003), with the finding that:
  - (a) both children and adults are sensitive to statistical regularities in artificial grammars
  - (b) artificial grammars are acquired with little or no explicit training
  - (c) with minimal training, learners are sensitive to phonological naturalness (Peperkamp Skoruppa and Dupoux, in press)
- Little research has been done on addressing the nature of *linguistic representations*, specifically, the nature of *phonological features*
- The nature of phonological representations can be addressed by testing the level of *generalization* that learners employ in artificial language tasks
  - *Generalization:* extension of a pattern to novel forms
    - *novel forms*: provides evidence of rule learning
    - *novel segments*: provides evidence of abstract representations
- Assessing the level of generalization makes it possible to understand the representations used in phonological learning

RESEARCH QUESTION

Do adult learners make use of feature or segment based representations in artificial language learning?

PRESENT EXPERIMENTS

To address this question, we make use of the Poverty of the Stimulus Paradigm (Wilson, in press):

-*Poverty of the Stimulus Paradigm*: Assesses level of representation of artificial language learning by exposing participants to a subset of the phonological inventory at training, *holding out* particular segments at test

- The present experiments follow Pycha et al. (2003) using a vowel harmony process as training

-*Vowel Harmony*: a phonological processes whereby vowels are forced to agree in some feature value (front, round, height, tenseness)
-occurs in wide variety of languages (Turkish, Finnish, Akan, Yokuts, etc.)

-Using front/back(Experiments 1 and 2) and height (Experiment 3) vowel harmony
- Expose participants to 4 vowels in a 6-vowel inventory for training
- Test using all 6 vowels
- Generalization to novel segments entails feature-based learning

-Three possibilities for learning a phonological process
- (a) *Segment-Based Learning Hypothesis*: learners form their rule based entirely on individual segments, and will not generalize to novel segments
- (b) *General Feature-Based Learning Hypothesis*: learners will form the most general possible rule to fit all of the data, and will generalize if possible.
- (c) *Restrictive Feature-Based Learning Hypothesis:* learners will form the most specific rule possible that will fit to the data and will generalize only to certain novel segments.(e.g. generalize only to [–Low] segments if only non-low data are given)

METHOD

Materials

Experiment 1: Front Harmony: High Suffix

- All suffixes alternated between [-mi] and [-mu]
- One condition had *mid* vowels held out, the other *low* vowels

	Front Vowels		Back Vowels	
Mid Hold-Out	[bimi]	[bimimi]	[tudu]	[tudumu]
	[pædi]	[pædimi]	[madu]	[madumu]
Low Hold-Out	[nege]	[negegemi]		[gomo] [gomomu]
	[degi]	[degimi]	[muto]	[mutomu]

Experiment 2: Front harmony, Low suffix vowel

-Suffixes alternated between [-mæk] and [-mak]

- One condition had *mid* vowels held out, the other *high* vowels

Experiment 3: Height Harmony

-All suffixes alternated between [-mi] and [-me]

- One condition had *lax* vowels held out, the other *back* vowels

-*Participants*

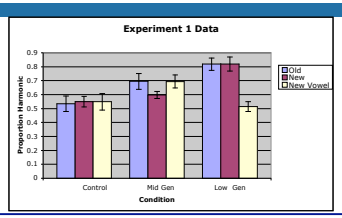
Johns Hopkins University students who received course credit for their participation.

Procedure

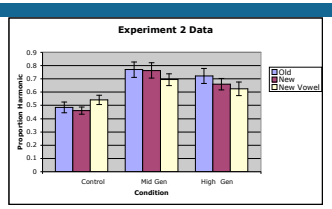
- Participants heard a randomized list of 24 stem , stem + suffix pairs 5 times
  -[bada, badamu], [mepe, mepemi], etc.
- Participants in control condition heard mixture of harmonic and disharmonic stems
- Forced-choice test: Harmonic vs. Disharmonic

[badamu] vs. *[badami]

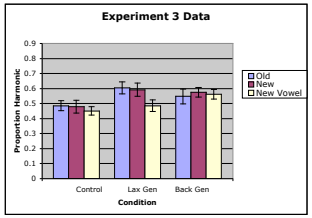
RESULTS — EXPERIMENT 1



RESULTS — EXPERIMENT 2



RESULTS — EXPERIMENT 3



CONCLUSIONS

Experiment 1 — Participants generalized to mid vowels, but not low vowels. This occurred even in the mid vowel condition, suggesting a general disfavoring of low vowels participating in the harmony rule. This may be due to a confound in the suffix alternation [-mi]/[-mu], which alternate in backness and rounding.

Experiment 2 — Participants generalized to both mid and high vowels, but less robustly for high vowels, specifically the high front vowel [i] reflecting the cross-linguistic tendency for high front vowels to be non-participating in backness vowel harmony.

Experiment 3 — Participants generalized to back vowels, but not lax vowels, reflecting the cross-linguistic tendency for height harmony systems to be dependent on tenseness but not backness

Overall Conclusions — The fact that participants generalized to segments in all three experiments supports the feature-based learning hypothesis. However, limits on these generalization reflect cross-linguistic tendencies in featural participation rather than formal restrictions on the specificity of feature values in rule formation

FUTURE RESEARCH

- Replicate the experiment using other phenomena besides vowel harmony
- Replicate the experiment in infants (see Seidl and Buckley, 2005 for infant studies using hold-out conditions)
- Using non-English populations
- Other types of harmony (tense/lax, round)
- Can participants generalize to novel vowels in the suffix? (are there trigger/target asymmetries in feature-based learning?)
- Experiments that further test the nature of feature-based learning, and the general psychological status of phonological representations

REFERENCES

Newport, E., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127-162.

Peperkamp, S., Skoruppa, K., & Dupoux, E. (in press). *The role of phonetic naturalness in phonological rule acquisition.* Paper presented at the Boston University Conference on Language Development.

Pycha, A., Nowak, P., Shin, E., & Shertzed, R. (2003). Phonological rule-learning and its implications for a theory of vowel harmony. *WCCFL, 22*, 101-113.

Seidl, A., & Buckley, E. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development, 1*, 289-316.

Wilson, C. (2003). Experimental investigations of phonological naturalness. *WCCFL, 22*, 101-114.

Wilson, C. (in press). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*.

Acknowledgements
Becky Piorkowsky, Rebecca Morley, Paul Smolensky, Luigi Burzio, Ariel Goldberg and the JHU Psycholinguistics Lab

*finley@cogsci.jhu.edu