

## Modularity bias in human and artificial learners

Elliott Moreton  
UNC-Chapel Hill

Phonological dependencies in natural language tend to relate elements which are phonetically similar. Two main factors may be at the root of this typological asymmetry. One is *channel bias*, errors in transmission between speaker and hearer which introduce phonetically-systematic biases into the corpus perceived by the learner (Ohala, 1994; Barnes, 2002; Blevins, 2004). The other is *analytic bias*, cognitive predispositions which render some patterns inherently harder to learn than others (Wilson, 2003). This talk presents evidence that there is a typologically effective analytic bias favoring dependencies between two instances of the same feature over dependencies between two different features, and shows that this bias can emerge in a learner without being explicitly hard-wired in.

### Learning experiments

Height-height and tone-tone dependencies are more frequent in natural language than height-voice or voice-tone dependencies, even though their phonetic precursors are of about the same size, and previous experiments have found that a height-height pattern is learned more readily in the lab than a height-voice one (Moreton, 2008). This suggests that the typological asymmetry is due to an analytic bias, but leaves several possibilities as to precisely what that bias favors:

- Dependencies which are typologically common.
- Dependencies between vowels.
- Dependencies which involve one feature.
- Dependencies on a single autosegmental or Feature-Geometric tier.
- Dependencies between featurally-similar elements.

A series of six minimally-different pattern-learning experiments with native English speakers supports the hypothesis that single-feature dependencies are detected more readily than two-feature dependencies.

### Modelling featural-simplicity bias

What could make a learner more receptive to single-feature than two-feature dependencies? One possibility is a detailed set of hard-wired learning biases. I will discuss an alternative in which bias emerges out of

- the learner's preference for "good explanations", i.e., hypotheses which make the training data relatively probable, and
- the lesser interaction between the critical dependency and the rest of the stimulus when the dependency involves two instances of the same feature.

The learner relies on the "Bayesian Occam's Razor" (MacKay, 2003, 343ff), rather than explicit penalties against particular pattern types. Simulation results provide a good qualitative (and fair quantitative) match to the human experimental data.

## References

- Barnes, J. (2002). *Positional neutralization: a phonologization approach to typological patterns*. Ph. D. thesis, University of California, Berkeley.
- Blevins, J. (2004). *Evolutionary phonology*. Cambridge: Cambridge University Press.
- MacKay, D. J. C. (2003). *Information theory, inference, and learning algorithms*. Cambridge, England: Cambridge University Press.
- Moreton, E. (2008). Analytic bias and phonological typology. *Phonology* 25(1), 83–127.
- Ohala, J. J. (1994). Towards a universal, phonetically-based theory of vowel harmony. In *Proceedings of the International Conference on Spoken Language Processing*, pp. 491–494.
- Wilson, C. (2003, January). Analytic bias in artificial phonology learning: consonant harmony vs. random alternation. Handout from presentation at the Workshop on Markedness and the Lexicon, Massachusetts Institute of Technology.