Parametric Linguistic Systems: The Limits of Probabilistic Learning for Realistic Data

Lisa Pearl (lpearl@uci.edu): University of California, Irvine

Mar 4-6, 2009

A strength of probabilistic learning is that it often allows children to make generalizations from ambiguous and noisy data effectively. Here, we focus on how young children could acquire a parametric system of metrical phonology (adapted from Dresher (1999) and Hayes (1995)), using English as a case study. We examine several cognitively plausible probabilistic learning models and find that, given empirical data as input, purely probabilistic models fail to converge on the English grammar reliably while models with a selective learning bias succeed.

The learning task for English metrical phonology is not trivial; the system explored here has 9 interacting parameters, making it difficult to determine which grammar generated a given data point. Moreover, the English data are highly noisy. Data from the Bernstein and Brent corpora in the CHILDES database (MacWhinney 2000) suggest that over 27% of the data are incompatible with the English grammar for at least one parameter value. Together, these factors make the learnability of the English metrical phonology system an interesting case study.

The unbiased probabilistic learning models examined here are adapted from the cognitively plausible model in Yang (2002), and assume children process data incrementally. Several variations of incremental learning models are tried, including those with batch learning and prior knowledge of part of the English system. Yet all these models fail. Close examination of the data reveals something that will trouble *any* unbiased probabilistic learner: the correct English grammar is not the one most compatible with the English child-directed speech data. So, a purely probabilistic learner would be unlikely to choose it.

Still, since English children seem to learn the English grammar, there must be some bias they bring to the learning scenario that makes the English grammar the most compatible. One bias that does lead to the correct learning behavior is to learn only from the subset of the available input that is perceived as *unambiguous* (Pearl 2008). When a learner uses this selective learning bias, the English grammar is learned from the English data.

While this may not be the only successful bias, the results here suggest an unbiased probabilistic learner will flounder when given realistic English data. Previous results for English provide one learning bias that will generate the correct learning behavior. Because the correct grammar *can* be learned -by biased learners- from the empirical data children use as input, this supports the parameter system as a viable hypothesis space instantiation that children could use.

Dresher, E. (1999). Charting the learning path: Cues to parameter setting. *Linguistic Inquiry*, *30*, 27-67.

Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk.* Mahwah, NJ: Lawrence Erlbaum Associates.

Pearl, L. (2008). Putting the Emphasis on Unambiguous: The Feasibility of Data Filtering for Learning English Metrical Phonology, in Chan, H., Jacob, H., and Kapia, E (eds.), *BUCLD 32: Proceedings of the 32nd annual Boston University Conference on Child Language Development,* Somerville, MA: Cascadilla Press, 390-401.

Yang, C. (2002). *Knowledge and Learning in Natural Language.* Oxford: Oxford University Press.