Constrained Probabilistic Learning for Complex Linguistic Systems

Lisa Pearl, UC Irvine
April 7, 2008
UCSD Linguistics Colloquium

Language learning is a tricky business. There are multiple kinds of knowledge a child must learn, some more transparently related to the observable linguistic data than others. In addition, the data are often noisy. Yet, despite these difficulties, children seem to always converge on the correct linguistic information for their native language. This talk focuses on how children acquire complex linguistic systems that are less easily inferable from the data, using the metrical phonology system as a case study.

Because human learning is gradual and reasonably robust to noise in the data, some kind of probabilistic learning is necessary. However, without constraints on what linguistic systems are possible, the hypothesis space for the learner is *infinite* – and even a probabilistic learner will have difficulty choosing the correct system in such a situation. Linguistic theory provides one idea for how to constrain the learner's hypothesis space: learners are guided by a selection of parameters (Chomsky 1981, Halle & Vergnaud 1987) or constraints (Tesar and Smolensky 2000) that enumerate the range of possible linguistic systems in human languages. The learner's task then is to converge on the correct linguistic system within this subset, using the data available from the native language. But the task is still not easy – the available data are often ambiguous and exception-filled. An interpretive bias to use only highly informative data (Fodor 1998, Dresher 1999, Lightfoot 1999, Pearl & Weinberg 2007) may help with this troublesome aspect.

This talk will explore the learnability of a parametric instantiation of the English metrical phonology system (Hayes 1995, Dresher 1999). The data is extrapolated from English child-directed speech (CHILDES: MacWhinney 2000), and is highly noisy. We will examine the performance of probabilistic learning algorithms that are psychologically plausible (adapted from Yang (2002)), and manipulate the kind of biases learners have. We will find that an interpretive bias is actually crucial for learning the correct system, even when the hypothesis space is tightly constrained. These results highlight the necessity of something beyond simple probabilistic learning – whether in the form of constraints on the hypothesis space or an interpretive bias for the data. In addition, because the parametric system is in fact learnable from realistic data, these results support the viability of hypothesis space restriction via linguistic parameters.

REFERENCES:
Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
Dresher, E. (1999). Charting the learning path: Cues to parameter setting. *Linguistic Inquiry*, *30*, 27-67.
Fodor, J. D. (1998). Unambiguous Triggers. *Linguistic Inquiry, 29*, 1-36.
Halle, M. & Vergnaud, J-R. (1987). *An essay on stress*. Cambridge, MA: The MIT Press.
Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
Lightfoot, D. (1999). *The Development of Language: Acquisition, Change, and Evolution*. Oxford: Blackwell.
MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Mahwah, NJ: Lawrence Erlbaum Associates.
Pearl, L., & Weinberg, A. (2007). Input Filtering in Syntactic Acquisition: Answers from Language Change Modeling. *Language Learning and Development, 3(1),* 43-72.
Tesar, B. & Smolensky, P. (2000). *Learnability in Optimality Theory*. Cambridge, MA: The MIT Press.
Yang, C. (2002). *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.