Linguistic Learnability, Fall 2021

William Sakas, sakas@hunter.cuny.edu

Every day, we use language so effortlessly that we often overlook its complexity. The fact that language *is* complex is indisputable. After decades of scrutiny, highly-trained adult scientists cannot agree on a definitive analysis of the underlying mechanisms that ultimately determine how our sounds, words, and sentences go together – but such an effortless task for a child! By the time a child reaches kindergarten, he or she has almost full mastery of an elaborate structure that eludes adequate scientific description. How children accomplish this – how they come to acquire 'knowledge' of language's essential organization – is one of the most fundamental, beguiling, and surprisingly open questions of modern science.

One line of investigation is to theoretically model the acquisition process from a mathematical and computational perspective in order to establish what bounds must exit on actual human language learning. This is the field of *Linguistic Learnability.* Learnability is a separate field from the study of *Language Acquisition* in that Learnability is focused on theoretical aspects of language acquisition, whereas Language Acquisition research explores empirical language interaction between children and their caregivers. For sure there is overlap. Equally true is that to be a well-rounded researcher in the acquisition of language by young children you need to be acquainted in both areas, even though your focus will likely be in one or the other. On the flip side, recent natural language processing (NLP) research has introduced techniques derived from the study of how humans actually process language and topics from this course could prove useful in future engineering endeavors.

The course will cover basic concepts such as poverty of the stimulus, the Subset Principle, parametric variation, 'triggering', defaults, error-driven learning, statistical learning, etc. Also, some new results on syntactic parameter setting from our CoLAG (Computational Language Acquisition) research group here at CUNY.

Topics:
1. How can kids possibly learn language?
   a. No negative evidence
   b. Poverty of the Stimulus
   c. Formal proofs: Gold and the Subset Principle
   d. Statistics, feasibility, "Ideal" learning
   e. Deterministic vs. trial-and-error learning
2. Parameter setting
   a. Gibson and Wexler's Triggering Learning Algorithm
   b. Yang's Variational Learner
   c. Fodor's Structural Triggers Learner
   d. Sakas and Fodor's Disambiguating Paradigm
   e. Recent work from CoLAG
3. Non-parameter setting
   a. N-gram modeling
   b. Artificial Neural Networks
   c. Bayesian modeling
   d. Information retrieval and other NLP tasks (overview if time allows)