Janet Dean Fodor
Graduate Center, CUNY

Evaluating Models of Parameter Setting

The CUNY Computational Language Acquisition Group (CUNY-CoLAG) consists of faculty and students at the CUNY Graduate Center. Names of all recent participants are below. An investment of effort over the last three years has begun to produce a wealth of data. Twelve different procedures for setting syntactic parameters have been programmed and put to the test in computer simulations. The input to these learning procedures is a stream of sentences drawn from a target language in a domain of several thousand languages each defined by universal principles with language-specific parameter values. The sentences have fully specified X-bar structure and movement transformations implemented as HPSG local dependencies. The test of a learner's efficiency is how many (or rather, how *few*) sentences it needs as input before arriving at the target grammar.

The computational and linguistic mechanisms underlying the construction of the domain will be illustrated in a demo of a pilot web interface. Data will then be presented from computer simulations on 12 parameter-setting algorithms and 4 target languages. (Since only small corners of the data have so far been collated and interpreted, we issue due warnings that final conclusions may differ from those to be presented.) It appears that there is a distinct advantage for learners that can "decode" the parameter values that are able to derive a given input sentence, i.e., the sentence structure guides the learner toward a viable set of parameter values. This contrasts with learning procedures which guess a grammar *before* finding out whether it is even capable of generating the sentence. Though obviously valuable, parametric decoding has not been a feature of all models of learning, because it has seemed to demand implausible computational resources. The *Structural Triggers* models developed at CUNY are at present the only ones that incorporate a psychologically plausible decoding system. Future developments will include tailoring the input stream to better match the distribution of sentence types and sentence complexity in real-life child-directed speech, as found in the CHILDES database.

Computer simulation: Carrie Crowther, Eiji Nishimoto, William Sakas, Virginia Teller
Language domain design: Janet Fodor, Iana Melnikova, Erika Troseth
CHILDES database research: Carrie Crowther, Atsu Inoue, Yukiko Koizumi, Lisa Reisig-Ferrazzano, Tanya Viger
Web design: Aaron Harnly