The IULC and the Linguistics Department are proud to present

John A. Goldsmith

The University of Chicago

Optimization is the answer.

Now, what is the question?

Friday, September 19
Ballantine Hall 103, 4:00

Keynote address of the IULC 40th Anniversary Colloquium Series

Abstract: My goal in this presentation is to offer a brief introduction to a view of linguistics which is empiricist, and which puts a heavy emphasis on the character of language learning, without being cognitivist. It is a view that says that the goal of the linguist is to understand how language *can* be learned, a goal distinct from that of the psychologist, who aims to understand (in a different sense) how language *is* learned.

On an empiricist account, the goal of the linguist is, first, to develop increasingly refined data regarding language use; second, to develop insightful and compact theories of the data; and third, to evaluate competing models with regard to two criteria: their abilities to *concisely* characterize regularities within and across languages, and their ability to identify *all* the generalizations that inhere in the data.

Probabilistic methods provide an explicit framework in which to accomplish such a task. Probability plays a role at two levels: at the lower, grammatical level, we place a condition on a grammar that it must assign a probability distribution over all the representations it generates (hence, the infinite sum of the probabilities must sum to 1.0). Secondly, probability plays a role at the higher, theoretical level, in that we must establish a "prior distribution over grammars"--which is to say, that a probability is assigned to the infinite class of grammars as well. This latter notion of probability is very similar to the classical generative notion of a simplicity metric (or its inverse, a complexity metric): the complexity of a grammar is closely related to the shorted possible length of the grammar expressed as a program on a universal computer.

For an empiricist account of linguistics, then, the *optimal* grammatical description of a finite set of data is that grammar which minimizes a quantity which is the sum of two terms: the length of the grammar, plus what is called the optimal compressed length of the data, given the grammar (this "optimal compressed length" of the data is equal to -log_2 of the probability of the data, given the grammar).

What does this mean for morphologists and phonologists? I will give four illustrations: (1) an account of word learning, (2) an account of morphological segmentation, (3) an account of sonority, and (4) an account of vowel harmony.