the error-driven ranking model of the acquisition of phonotactics
Giorgio Magri
(Institut Jean Nicod)
Paris, 24 novembre 2011

Children learn their target language early, fast and efficiently. For instance, nine-month-old infants already display knowledge of the phonotactics of their target language, namely have been shown to react differently to licit versus illicit sound combinations. Children must thus rely on a remarkably efficient phonotactic learning strategy. What could it look like? According to the error-driven learning model, the learner maintains a current hypothesis of the target adult phonotactics and keeps slightly updating its current hypothesis whenever it makes a mistake on the incoming stream of data from the adult language, until it makes no more mistakes. This learning model is very popular in the current Optimality Theoretic (OT) acquisition literature because of its cognitive plausibility: it models the observed acquisition gradualness, as it describes a stepwise progression towards the target adult grammar; it relies on surface phonology without requiring any knowledge of morphology, that plausibly develops later than phonotactics; and it does not impose unrealistic memory requirements, as it only looks at a piece of data at the time without keeping track of previously seen data. Yet, the OT computational literature has failed so far to develop a computationally sound implementation of OT error-driven learning. In particular, Tesar & Smolensky's implementation is too simple to match the observed acquisition complexity; and Boersma's implementation is not computationally sound. My current project aims at filling this gap, thus establishing OT error-driven learning as a sound and plausible model of the acquisition of phonotactics. Combining a computational perspective focused on learnability with a modeling perspective based on large databases of child acquisition data, the project addresses five core issues. The first issue concerns convergence: does the model eventually stop making mistakes and settle on a final grammar? The second issue concerns correctness: does the final grammar entertained by the model indeed capture the target phonotactics? The third issue concerns modeling adequacy: do the learning sequences formally predicted by the model match attested child acquisition paths? The fourth issue concerns robustness and variation: how does the model behave in the presence of noise and how can it make sense of the pervasive phenomenon of child variation? The fifth issue concerns framework selection: how can the choice of the OT framework be justified from a learning theoretic perspective? This talk will provide an overview of the project, with a focus on some recent results concerning the first two issues of convergence and correctness.

La présentation aura lieu dans le cadre du **LingLunch Paris Diderot**, organisé chaque jeudi à l'UFR de Linguistique de l'Université Paris Diderot.
Salle 4C92, 12:00 à 13:00; 175 rue du Chevaleret, 75013 Paris, M° Chevaleret.
    http://www.linguist.univ-paris-diderot.fr/linglunch.html