

Bootstrapping into phonology

Diana Apoussidou
University of Amsterdam

This paper presents an account for the learning of underlying forms in Optimality Theory (OT; Prince & Smolensky 1993). Accounts in OT for the determination of underlying forms, such as Lexicon Optimization (Prince & Smolensky 1993, Itô et al. 1995), compare different forms of a paradigm and usually presuppose the ranking of the language. This is not applicable to a realistic learning situation. A child acquiring a language does neither know the ranking beforehand nor the underlying forms, and it does not have access to all the different forms of a paradigm at once. I therefore propose a learnability account for the concurrent acquisition of the grammar and the underlying forms in the lexicon. The assumed model has four levels of representation with one phonetic representation (the *overt form* OF which is directly accessible to a learner), two phonological representations (the *surface form* SF as the grammatical phonological form and the *underlying form* UF as the lexical phonological form) and one semantic representation (the *meaning* M of a form). The different representations are connected via different constraint families: OF is connected to SF via structural constraints; SF is connected to UF via faithfulness constraints; and UF is connected to M via lexical constraints. I argue that the concurrent learning of both SF and UF is possible if the language learner is provided with OF and its corresponding M. With these two representations at hand, the learner can bootstrap into phonology (SF and UF).

I exemplify the proposal with word stress in Modern Greek. Stress in this language is mainly lexical, i.e. the underlying morphemes of a word can be specified in the lexicon for stress, and there are different kinds of specifications. For instance, *γόνδολα* ‘gondola-Nom.Sg.’ and *θάλασα* ‘sea-Nom.Sg.’ look the same in the nominative case, but when they are inflected with genitive plural case, they behave differently: *γόνδολα* becomes *γόνδολον*, with stress on the same position, but *θάλασα* becomes *θαλασόν*, with stress on the final syllable. Another case is *καλόγηρος* ‘monk-Nom.Sg.’ and *άνθρωπος* ‘man-Nom.Sg.’. When inflected with the genitive singular suffix *-u*, *καλόγηρος* becomes *καλόγηρου* (stress does not shift), while *άνθρωπος* becomes *ανθρώπου* (stress shifts). It has been argued (e.g. by Revithiadou 1999) that in a word like *γόνδολα*, the root *γόνδολ-* is underlyingly stressed, and the root *θαλασ-* is not. In combination with an underlyingly unstressed suffix, like *-a*, the word is assigned a default stress, but in combination with an underlyingly stressed suffix like *-on*, the suffix can assign stress to the word. In the case of *καλόγηρου* and *ανθρώπου*, the suffix *-u* is specified for being pre-stressing, and can assign stress to the word when combined with an unmarked root like *άνθρωπ-*. A child acquiring Greek needs to learn which morpheme has an underlying specification for stress, and what kind of specification.

I will support my claim with computer simulations of the acquisition of underlying forms. The ingredients are: 1. a constraint set accounting for the stress pattern in Greek (structural, faithfulness, and lexical constraints); 2. a learning algorithm for OT grammars, the Gradual Learning Algorithm (Boersma 1997); and 3. learning data in the form of pairs of overt form and meaning (i.e. Greek words with different stress patterns).

The results show that the learning of underlying forms is possible with this on-line approach. However, it appears that the computer-simulated learners can have variation on the two phonological levels SF and UF. This variation is not apparent (i.e. not observable in the phonetic signal), since OF does not show variation. The proposed model furthermore implies a partial grammaticalization of the lexicon, since underlying forms are regulated by lexical constraints.

References

- Boersma, P. (1997). How we learn variation, optionality, and probability. *IFA Proceedings* 21:43-58.
- Itô, J., A. Mester & J. Padgett (1995). Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26:571-613.
- Prince, A. & P. Smolensky (1993). Optimality Theory: constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick, NJ, and University of Colorado, Boulder. Technical Report RuCCS-TR-2. Rutgers Center for Cognitive Science.
- Revithiadou, A. (1998). *Headmost accent wins*. Doctoral dissertation, LOT.