Reducing linguistic variation to Third Factor mechanisms

Jordi Fortuny (UB, CLT-UAB, Complex Systems Lab-UPF) & Adriana Fasanella (CLT-UAB)

The objective of our study is to develop a model of morphophonological analysis that enables the learner to infer high-order properties of the target language. Our first step is to express in a parametric way part of the mechanism of data analysis used by the Language Acquisition Device (LAD) in order to attain a morphological analysis of its Primary Linguistic Data (PLD). The second step is to explore how these analyses can be used by the learner to deduce classical patterns of morphosyntactic variation. With this approach we show a path to reduce the problem of what is the permissible format of language variation to the Third Factor mechanisms (Chomsky 2005) responsible for language acquisition.

The starting intuition is that all languages share the same class of grammatical features but differ as to how they realize them morphophonologically (Cinque 1999). We consider the minimal morphological category, which we will call *morph* or *head*, as a primitive of the proposed procedure, which can be detected on the PLD:

Definition. A linguistic form α, viewed as a string of phonemes, is a morph or head iff it is meaningful and does not contain any meaningful non-empty proper substring.

The properties to be set by the mechanism of data analysis under consideration are the following:

1) A head is *bound* if it is phonologically dependent of other heads and *unbound* otherwise.

2) A head is *synthetic* if it conveys more than one morpheme and *non-synthetic* if it conveys only one morpheme.

Property 1) is fixed by the learner by inspecting the string of heads. Whether a head is bound or not is arguably determined on the basis of phonological cues in the acoustic signal, such as pauses. Language-specific cues may also play a role, such as word level stress patterns, phonotactic regularities and allophonic variation. Property 2) is fixed by inspecting how a head is related to grammatical categories provided by Universal Grammar (UG), henceforth *morphemes*. More precisely, the mechanism should inspect how a head is related to morphemes, whether it conveys a sole morpheme or more. Here not only mechanisms of speech segmentation are involved, but the set of grammatical categories provided by UG and a theory of paradigmatic relations (Pinker 1984) must also be taken into consideration.

We call the morphophonological analysis mechanism we want to explore *Chunking Procedure*, and we understood it as follows:

3) **Chunking Procedure**. Given a head *H*, the learner determines whether *H* is phonologically dependent of other heads ([+bound]) or not ([-bound]); and whether *H* conveys only one morpheme ([-synthetic]) or more ([+synthetic]).

Once this morphological analysis is attained, we investigate the existence of bootstrapping mechanisms that use its results to specify higher order syntactic properties of the target language, namely those properties that traditional parameters range over. We capitalize on the observation that there exist general correlations between abstract syntactic patterns and the morphophonological analysis obtained by the Chunking Procedure. We shall directly formulate these correlations as bootstrapping mechanisms:

4) **Bootstrapping mechanisms tiggered by the Chunking Procedure**

(a) Once the learner has determined that there is a [+bound] head instantiating a feature *F*, then he can infer that the maximal projection instantiating *F* in the target language has a free distribution, and can be omitted.

(b) Once the learner has determined that there is a [+bound] head conveying *case* or *number* on pronouns, then he can infer that any argument of the verb can be omitted in the target language.

(c) Once the learner has determined that there is a [-bound] or a [+bound, -synthetic] head expressing *path*, then he can infer that multiple constructions that are related with the separate lexicalization of this head are available in the target language.

We shall sketch how the Chunking Procedure may be used to shed light on the problem of how the LAD infers syntactic properties of the target language from a morphophonological analysis in three selected case studies.

I. *Baker's (1996) Polysynthesis Parameter*. Assume that, given an amount of linguistic input, the Chunking Procedure has determined that there is a [+bound] head H_1 that instantiates a particular θ-role θ_1. The LAD should be able to determine on independent grounds whether H_1 is an incorporated noun or an affix agreing with a DP; if H_1 can also appear without being incorporated and as a fragment, then it will be a noun, whereas if H_1 is always bound (i.e., it cannot appear freely or as a fragment), then it will be an affix. Consider now the latter situation, in which H_1 is an affix agreeing with a maximal projection. In virtue of the bootstrapping mechanism (4.a), it follows that the maximal projection which the affix agrees with can be omitted and has a relatively free distribution.

II. *Neeleman & Szendrői (2007)'s strong prediction on radical pro-drop*. Assume the LAD has detected in the linguistic input that there is a head H_1 instantiating the category of *case* or *number* analyzed as [+bound] with respect to pronouns. At this moment, the LAD follows the bootstrapping mechanism formulated in (4.b) and infers that the target language allows radical pro-drop, in which case verbal arguments and possessors can be omitted.

III. *Satellite-framed languages and related constructions (Talmy 1985)*. Assume that the Chunking Procedure has detected a H_1 expressing solely *path*; then there are two subcases: H_1 is [-bound] if the target language is a strong satellite-framed language, like English, or H_1 is [+bound, -synthetic] if the target language is a weak satellite-framed language, like Latin. In both cases, given the bootstrapping mechanism defined in (4.c), the LAD infers the availability of the relevant set of constructions (complex directed motions, unselected objects, complex effected objects, etc.).

Our approach consists, therefore, in coding parameters in mechanisms of morphological data analysis and deriving syntactic variation from the value attained by those mechanisms. This move suggest that Greenberg's problem (what the nature and format of permissible linguistic variation is) may be reduced to Plato's problem (how natural languages are learned). By using this methodology, linguistic variation is examined in the very same terms as those used by the LAD when analyzing the PLD and, consequently, morphosyntactic variation is constrained by mechanisms of data analysis active during the process of language acquisition. Furthermore, provided that procedures of data analysis are considered to be elements of Third Factor, this proposal leads to the appealing conclusion that by defining data analyzers in a parametric fashion, linguistic variation could be embodied in certain Third Factor mechanisms.

References

–Baker, M. 1996. *The polysynthesis Parameter* (Oxford: Oxford University Press).
–Chomsky, N. 2005. Three factors in language design, *Linguistic Inquiry* 36(1), 1-22.
–Cinque, G. 1999. *Adverbs and functional heads*: a cross-linguistic perspective (Oxford: Oxford University Press).
–Neeleman, A. & K. Szendrői. 2007. Radical pro-drop and the morphology of pronouns, *Linguistic Inquiry* 38(4), 671-714.
–Talmy, L. 1985. Lexicalization patterns: Semantic structure in lexical forms, *Language typology and syntactic description* 3, 57-149.