Empiricist Solutions to Nativist Puzzles

Rens Bod (rens.bod@gmail.com) Gideon Borensztajn (gideonbor@gmail.com)
Emily Morgan (eipmorgan@gmail.com)
Cognitive Science Center Amsterdam, Institute for Logic, Language and Computation
University of Amsterdam, Netherlands

Introduction

How much knowledge of language is innate and how much is learned through experience? The *nativist* view endorses that human language acquisition is guided by innate rules ("Universal Grammar"), while the *empiricist* view assumes that language acquisition is the product of abstractions from stored exemplars. Despite the apparent opposition between these views, the essence of the debate lies in the relative contribution of prior knowledge and linguistic experience. The main goal is then to establish the minimal prior knowledge needed for language acquisition to take place.

In this paper we will focus on a couple of hard linguistic phenomena that have for a long time been considered "parade cases of an innate constraint": *auxiliary fronting* and *binding*. We show that these linguistic facets can be learned without assuming specific constraints or principles and without presuming that they have been literally seen in the data. Instead, we will demonstrate that these phenomena can be learned either from *simpler* sentences in a corpus of child-directed speech (Bod 2009) or emerge from a specific parsing strategy.

Auxiliary Fronting (AF)

Based on Chomsky (1968), who argued that AF was a case for the Poverty of the Stimulus argument, Crain & Nakayama (1987) gave evidence that children never produce sentences of type (2). Moreover, they maintained that sentences of type (1) are so rare that children must have innately specified knowledge that allows them to learn AF without ever having heard it.

(1) Is the man who is eating hungry?
*(2) *Is the man who eating is hungry?*

However, Reali & Christiansen (2005) showed that AF can also be learned by linear models such as SRNs and *n*-grams. Yet, some of their success depends on "accidental" English facts (Kam et al. 2008). Clark and Eyraud (2007) showed that AF can be learned by inferring a CFG from distributional statistics, but Berwick & Chomsky (2008) opposed that their result is limited to an *artificial* example. In Bod (2009), we replicated the induction of AF, and additionally showed that AF can be learned from *realistic* child-directed speech in the Eve (Childes) corpus. Although complex auxiliary fronting is not literally recorded in the Eve corpus, it can be 'elicited' by combining chunks from previous utterances by means of simple substitution in the shortest possible way. It turns out that this model can also simulate various AF-errors made by children.

Binding and C-command

Another widely studied linguistic phenomenon is binding, wherein multiple noun phrases are understood to have the same referent. Crain and Thornton (2006) motivate a discussion of binding with the following sentences:

(3) While the Ninja Turtle$_i$ danced, he$_i$ ate pizza.
(4) While he$_i$ ate pizza the Ninja Turtle$_i$ danced.
(5) The Ninja Turtle$_i$ danced while he$_i$ ate pizza.
*(6) *He$_i$ ate pizza while the Ninja Turtle$_i$ danced.*

In examples (3-5), *he* can be coreferential with *the Ninja Turtle,* but in (6) it cannot. This fact is traditionally explained in terms of a constraint known as *Binding Principle C*, which in turn relies on one of the most fundamental notions in generative grammar, known as *c-command*. It turns out that Principle C is violated in (6) while not in examples (3-5). However, we will show that the binding constraint can also emerge as a result of incremental parsing. There is no need for separate principles. Rather than viewing c-command as an explicit rule that can either be learned or innately granted, we show that it becomes a natural relation when using an incremental left-corner parser. This suggests that other 'hard' phenomena, which also rely on c-command, can be explained in an empiricist way as well.

References

Berwick, R. & Chomsky, N. 'Poverty of the Stimulus' revisited: recent challenges reconsidered. In *Proceedings CogSci 2008*, 383.

Bod, R. (1998). *Beyond Grammar.* Stanford: CSLI Publications.

Bod, R. (2002). A Unified Model of Structural Organization in Language and Music. *JAIR*, *17*, 289-308.

Bod, R., Hay, J. & Jannedy, S. (eds.) (2003). *Probabilistic Linguistics*. The MIT Press.

Bod, R. (2006). Exemplar-Based Syntax: How to Get Productivity from Examples. *The Linguistic Review 23*, 291-320.

Bod, R. (2007). Is the End of Supervised Parsing in Sight? In *Proceedings ACL 2007,* 400-407.

Bod, R. (2009). From exemplar to grammar: a probabilistic analogy-based model of language learning. *Cognitive Science* 33, In press. <http://staff.science.uva.nl/~rens/analogy.pdf>.

Chomksy, N. (1968). *Language and Mind*. New York: Harcourt.

Clark, A. & Eyraud, R. (2007). Polynomial Identification in the Limit of Substitutable Context-free Languages, *Journal of Machine Learning Research* , 8, 1725-1745.

Crain, S., & Nakayama, M. (1987). Structure Dependence in Grammar Formation. *Language, 63*, 522-543.

Crain, S., & Thornton, R. (2006). Acquisition of syntax and semantics. In M. Thraxler & M. Gernsbacher (eds.), *Handbook of Pscyholinguistics*, Elsevier, 1073-1110.

Kam, X.-N, Stoyneshka, I., Tornyova, L., Fodor, J., & Sakas, W. (2008). Bigrams and the Richness of the Stimulus. *Cognitive Science*, *32*, 771-787.

Reali, F., & Christiansen, M. (2005). Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence. *Cognitive Science*, *29*, 1007-1028.