In Piatelli et al, M. *Of minds and language: the Basque country encounter with Noam Chomsky.* Oxford University Press. 279-295 2009

CHAPTER 18

Remarks on the Individual Basis for Linguistic Structures*

Thomas G. Bever

This paper reviews an approach to the enterprise of paring away universals of attested languages to reveal the essential universals that require their own explanation. An example, discussed at this conference, is the long-standing puzzle presented by the Extended Projection Principle (EPP, Chomsky 1981). I am suggesting an explanation for the EPP based on the learner's need for constructions to have a common superficial form, with common thematic relations, the hallmark of EPP. If one treats EPP phenomena as the result of normal processes of language acquisition, the phenomena not only receive an independently motivated explanation, they also no longer constitute a structural anomaly in syntactic theory.[1]

The concept of "language" is like those of ... "organ", as used in biological science ... grammatical structure "is" the language only given the child's intellectual environment ... and the processes of physiological and cognitive development ... Our first task in the study of a particular [linguistic] structure in adult language behavior is to ascertain its source rather than immediately assuming that it is *grammatically* relevant ... Many an aspect of adult ... linguistic structure is itself partially determined by the learning and behavioral processes that are involved in acquiring and implementing that structure ... Thus, some formally possible structures will never appear in any language because no child can use [or learn] them. (Bever 1970: 279–280)[2]

Here I focus on the dynamic role of the individual language learner in shaping properties of attested languages (aka E-languages). Certain linguistic universals that seem to be structural are in fact emergent properties of the interaction of genetic endowment, social context, and individual learning dynamics. My argument is this: Language acquisition recruits general mechanisms of growth, learning, and behavior in individual children: only those languages that comport with these mechanisms will be learned. I first review some non-syntactic universals, to outline relatively clear examples of the role of development, as background for the main focus of this paper.

18.8 Coda: Some broader implications of the AxS acquisition model

The following points are in large part the result of email discussions with Noam.

18.8.1 Language acquisition as enjoyable problem solving

The idea that the child acquires knowledge of syntax by way of compiling statistical generalizations and then analyzing them with its available syntactic capacities is but another instance of learning by hypothesis-testing. For example, it is technically an expansion on the TOTE model proposed by Miller et al. (1960). An initial condition (statistically grounded pattern) triggers a TEST meaning, and an OPERATION (derivation) which triggers a new TEST meaning and then EXIT. Karmilov-Smith and Inhelder (1973) advanced a different version – cognition advances in spurts, triggered by exposure to critical instances which violate an otherwise supported generalization.

The dual nature of the acquisition process is also related to classical theories of problem solving (e.g., Wertheimer 1925, 1945). On such models, the initial stage of problem organization involves noting a conceptual conflict – for example, "find a solution that includes both X and Y: if the answer is X then Y is impossible, but if Y then X is impossible": characteristically the solution involves accessing a different form of representation which expresses the relation between X and Y in more abstract terms. In language the initial conflict expresses itself as the superficial identity of all the constructions in (12) which exhibit the canonical form constraint, while assigning different semantic relations; the resolution is to find a derivational structure for the set that shows how the different surface constructions are both differentiated and related derivationally. Hence, not only is language-learning hereby interpreted in the context of a general set of learning principles, it is also interpreted as a special instance of a general problem solver. This also explains why language learning is fun, and hence intrinsically motivating: the gestalt-based model suggests that language-learning children can enjoy the "aha" insight experience, an intrinsically enjoyable sensation which may provide critical motivation to learn the derivational intricacies of language (cf. Weir's 1962 demonstration that children play with their language paradigms when they are alone).

Note that the terms "motivation" and "fun" are technical terms based in aesthetic theory, not the everyday notion of conscious desire, nor any notion of "reinforcement." Elsewhere, I have developed analyses of what makes objects and activities intrinsically enjoyable (Bever 1987). The analysis draws on the classic aesthetic definition: stimulation of a representational conflict which is then resolved by accessing a different form or level of knowledge. The formal similarity of this definition to the gestalt model of learning affords an explanation of why aesthetic objects are enjoyable: they are mini-"problems" involving conflicting representational solutions, resolved by accessing a level which creates a productive relation between those solutions, thereby eliciting a subconscious "aha." This kind of analysis is ordinarily applied to serial arts such as drama or music, in which the representational conflict and its resolution can be made explicit over time. But the analysis works for static objects, explaining the preference, for example, for the golden mean rectangle. In language, one kind of conflict is elicited by the thematic heterogeneity of superficially identical surface phrase structures: the child's resolution of that conflict requires access to an inner form of the sentences, via distinct derivational histories – a resolution which involves accessing a distinct level of representation. Thus, learning the structure of a language elicits a series of mini-ahas in the child, making it an activity which is intrinsically attractive.

The model also offers a partial answer to the frame problem (see Ford and Hayes 1993), the problem of how statistical generalizations are chosen out of the multiple possibilities afforded by any particular set of experiences. This problem was classically addressed by Peirce (1957) as the problem of abduction, who argued that there must be constraints on all kinds of hypotheses, even those ostensibly based on compilation of observations (cf. Chomsky 1959c, on the corresponding problem in S-R associative theory, and this volume). But the problem is also a moving target for the language-learning child. At any given age, the generalizations that are relevant to progress in learning are different: if the child has mastered simple declarative constructions, or some subpart of her language's inflectional system, this changes the import of further exposure to the language. Thus, we must not only address constraints on the initial state of the child (see Mehler and Bever 1968 for discussion), we must address how constraints apply to each current state of knowledge, as the child matures and acquires more structural knowledge. That is, the abductive constraints themselves have a developmental course. By what process and dynamic? Another way of putting this is, what filters (aka "frames") possible generalizations and how does the filter itself change as a function of current knowledge?

In the AxS scheme, there are two kinds of processes which filter generalizations. First is the set of salient regularities among elements that are available to the input: at a phonological level, infants have available perceptual categories that provide an initial organization of the input; this affords an innate categorization of sound sequences, available for formal derivational analysis. The other side of the filtering process is the set of computational devices available to provide a derivation. That is, those generalizations about sound sequences that endure are just those that can be explained by a set of possible computational phonological rules. Such rules must have natural domains (presumably innately determined) such as segmental features, syllabic structures, lexical templates. At the syntactic level, the corresponding problem is to isolate a natural segmentation of the potential compositional input. To put it in terms of the example we are focusing on, how does the system isolate "NP V (NP)" as a relevant kind of sequence over which to form a generalization? In the model proposed, the solution lies in the fact that the derivational component has its own natural units, namely clause-level computations. The result is that the derivational discovery component acts as a filter on the multiple possible statistical generalizations supported by any finite data set, picking out those that fit the derivational templates. Most important is that the properties of the derivational filter change as the knowledge base increases in refinement.