ACKNOWLEDGMENTS

This work was supported by a Fulbright research grant (No. 1202204) given to the author by CIES in the United States and the Hungarian-American Commission for Educational Exchange. Special thanks to Robert M. Harnish of the University of Arizona, Tucson, for useful suggestions in preparing the manuscript.

brain, Meaning, Grammar, evolution

Michael A. Arbib

Computer Science Department, Neuroscience Program, and USC Brain Project, University of Southern California, Los Angeles, CA 90089-2520.
arbib@pollux.usc.edu http://www-hbp.usc.edu/

Abstract: I reject Jackendoff's view of Universal Grammar as something that evolved biologically but applaud his integration of blackboard architectures. I thus recall the HEARSAY speech understanding system - the AI system that introduced the concept of "blackboard" - to provide another perspective on Jackendoff's architecture.

The subtitle "Brain, Meaning, Grammar, Evolution" for *Foundations of Language* (Jackendoff 2002) suggested that Jackendoff would devote major portions of his book to brain and evolution. Alas, there is no serious discussion of the brain (beyond a few passing references to aphasia) and the discussion of evolution (Ch. 8) focuses on an incremental account of Universal Grammar (UG) that ignores brain evolution. Space does not permit proper discussion of the brain here. Instead, I lament Jackendoff's view of Universal Grammar as something that evolved biologically; and then recall the HEARSAY speech understanding system to provide another perspective on Jackendoff's architecture.

Concerns about Universal Grammar. Jackendoff (2002, p. 263) views UG as "the unlearned basis from which language is learned" and argues that "it had better be available to help children learn case systems, agreement systems, fixed word order, and grammatical functions in case the language in the environment happens to have them."

I find this view incoherent if it implies that evolution yielded adaptations specific to each of these systems. What selective pressure would cause humans whose language does not use cases to evolve a brain with a device *specialized* for learning case systems?! Instead, I think we should seek to understand what made the brain "language ready," providing capacities that make possible the discovery of Jackendoff's language "components" over the course of many millennia, and their acquisition by the child over the course of a few years. One listing of such capacities (based on Arbib 2002b) follows:

Complex imitation: the ability to recognize another's performance as a combination of familiar movements and then repeat it.

Symbolization: The ability to associate an arbitrary symbol with a class of episodes, objects or actions. (These symbols may have been unitary utterances, rather than words in the modern sense, and may have been based on manual and facial gestures rather than being vocalized.)

Parity (mirror property): What counts for the "speaker" must count for the "listener."

Intentional communication: Communication is intended by the utterer to have a particular effect on the recipient, rather than being involuntary or a side effect of praxis.

From hierarchical structuring to temporal ordering: Perceiving that objects and actions have sub-parts; finding the appropriate timing of actions to achieve goals in relation to those hierarchically structured objects.

Beyond the here-and-now: The ability to recall past events or imagine future ones.

Paedomorphy and sociality: A prolonged period of infant dependency combines with social structures for caregiving to provide the conditions for complex social learning.

In hindsight we may see these as preadaptations for language but they were adaptive in their own right, and underlie many modern human capacities other than language. In this view, Universal Grammar is only tenable as a descriptive umbrella for the immense variety of human languages, not as a "genetic reality" or "neural reality" that implausibly contains all possible grammatical structures in embryo (one is reminded of the "little man" that seventeenth century spermists "saw" inside the head of the spermatozoon [Pinto-Correia 1996; 1997]). I applaud Jackendoff's attempt to provide an evolutionary sequence for language but argue (e.g., Arbib 2002b) that case systems, agreement systems, and so on, are to be seen as human inventions that required no change in brain structure for their discovery and cultural transmission. Moreover, I see these as coarse grain compared to the actual inventions that were made across the millennia and which eventually coalesced into the more-or-less coherent structures that Jackendoff and other linguists tend to treat as natural and indivisible. What is universal is the need for expression, not the choice of linguistic structure for meeting those needs. The evolution of language from protolanguage is part of the history, not the biology, of *Homo sapiens.*

Déjà-entendu. Jackendoff makes much of the AI notion of blackboard in presenting his architecture for language, but does not cite HEARSAY-II (Erman et al. 1980; Lesser et al. 1975), perhaps the first AI system to develop a blackboard architecture. While obviously not the state of the art, it is of interest because it foreshadows features of Jackendoff's architecture. Digitized speech data provide input at the *parameter level;* the output at the *phrasal level* interprets the speech signal as a sequence of words with associated syntactic and semantic structure. Because of ambiguities in the spoken input, a variety of hypotheses must be considered. To keep track of all these hypotheses, HEARSAY uses a dynamic global data structure, called the *blackboard,* partitioned into various levels; processes called *knowledge sources* act upon hypotheses at one level to generate hypotheses at another.

First, a knowledge source takes data from the *parameter level* to hypothesize a phoneme at the *surface-phonemic level.* Many different phonemes may be posted as possible interpretations of the same speech segment. A lexical knowledge source takes phoneme hypotheses and finds words in its dictionary that are consistent with the phoneme data - thus posting hypotheses at the *lexical level* and allowing certain phoneme hypotheses to be discarded. To obtain hypotheses at the *phrasal level,* knowledge sources embodying syntax and semantics are brought to bear. Each hypothesis is annotated with a number expressing the current confidence level assigned to it. Each hypothesis is explicitly linked to those it supports at another level. Knowledge sources cooperate and compete to limit ambiguities. In addition to data-driven processing which works upward, HEARSAY also uses hypothesis-driven processing so that when a hypothesis is formed on the basis of partial data, a search may be initiated to find supporting data at lower levels. A hypothesis activated with sufficient confidence will provide the context for determination of other hypotheses. However, such an *island of reliability* need not survive into the final interpretation of the sentence. All we can ask is that it forwards the process which eventually yields this interpretation.

Hanson and Riseman (1987) based the architecture of their computer vision system VISIONS on the HEARSAY architecture as well as neurally inspired schema theory (Arbib 1981; Arbib et al. 1998). Such a conceptual rapprochement between visual perception and speech understanding offers a computational framework for further exploration of the Saussurean sign (Arbib 2003; Hurford 2003). Arbib and Caplan (1979) discussed how the knowledge sources of HEARSAY, which were scheduled serially, might be replaced by schemas distributed across the brain to capture the spirit of "distributed localization" of Luria (e.g., Luria 1973). Today, advances in the understanding of distributed computation and the flood of brain imaging data make the time ripe for a new push at a neurolinguistics informed by the understanding of distributed computation. Despite its disappointing inattention to the brain, Jackendoff's book could make a valuable contri-

bution to this effort by showing generative linguists how to break out of the straitjacket of syntactocentrism by integrating their work into a rich multi-modal architecture.

ACKNOWLEDGMENT

Preparation of this commentary was supported in part by a Fellowship from the Center for Interdisciplinary Research of the University of Southern California.

Language evolution without evolution

Derek Bickerton

Department of Linguistics, University of Hawaii, Honolulu, HI 96822.
bickertond@prodigy.net

Abstract: Jackendoff's major syntactic exemplar is deeply unrepresentative of most syntactic relations and operations. His treatment of language evolution is vulnerable to Occam's Razor, hypothesizing stages of dubious independence and unexplained adaptiveness, and effectively divorcing the evolution of language from other aspects of human evolution. In particular, it ignores connections between language and the massive discontinuities in human cognitive evolution.

I approach Jackendoff's ambitious and many-faceted *Foundations of Language: Brain, Meaning, Grammar, Evolution* (Jackendoff 2002) as an unashamed syntactocentrist. Jackendoff, however, is far from being that, and the main example he picks to illustrate syntactic relations could hardly have been better chosen had he deliberately intended to marginalize and trivialize syntax:

(1) *The little star's beside the big star.*

This sentence, first analyzed on pages 5 through 6, is returned to repeatedly throughout the text.

But, copular sentences like (1), sentences with the verb "to be," form a small and highly idiosyncratic subset of sentences; their properties differ sharply from those of the vast majority of sentences. The latter describe actions, events, or a variety of states, and deploy a rich variety of argument structures; copular sentences express only identity, location, or the attribution of qualities (The rose is red/in the vase/a Molly Perkins) and take only a *theme* argument. In a non-copular clause, no two noun-phrases will have the same referent (unless a specifically reflexive form such as himself is used), and transposition of noun-phrases inevitably changes meaning:

(2) a. *John hit the captain.*
b. *The captain hit John.*

In copular clauses, no two noun-phrases will have different referents; consequently, transposition of noun-phrases inevitably leaves meaning unchanged:

(3) a. *John is the captain.*
b. *The captain is John.*

There are many more syntactic relations that can't be illustrated via copular sentences, too many to list here. Perhaps in his response to commentary Jackendoff will tell us why he chose such an atypical sentence as his prime syntactic exemplar.

Much more could be said about Jackendoff's treatment of syntax, but I must reserve the bulk of this commentary for his chapter on language evolution. Right off, Jackendoff confuses the issues with a straw-man version of "the common view of Universal Grammar" (p. 233). According to him, that view treats phonology and syntax as "passive handmaidens of syntax" that could not, therefore, have evolved prior to syntax. But syntax without phonology and semantics would be useless, so this view is absurd.

In fact the current status of semantics and phonology (whatever that may be) carries no entailment for their order of evolution. No one disputes that apes and hominids had some sort of conceptual structure, therefore semantics (in some form) had to precede syntax (indeed, this is made quite explicit in my own writings, from Bickerton 1990 on). As for phonology, this (at least in some primitive form) was presumably present in protolanguage, which had no syntax. But the emergence of syntax selected for a sophisticated phonology, while the capacity to assemble semantic units into complex propositions radically expanded conceptual structure.

Jackendoff then turns to the proposal of Bickerton (1990) that language developed in two steps, an asyntactic protolanguage and syntacticized modern language, and instead opts for "a more graceful, incremental evolution" (p. 236). But are the incremental stages he proposes really stages at all?

Take the three stages: (1) "use of symbols in a non-situation-specific fashion," (2) "use of an open, unlimited class of symbols," and (3) "development of a phonological combinatorial system" that supposedly intervene between an alingual state and protolanguage. No real difference exists between the first two. A symbol freed from the here and now has to be cultural rather than biological; if you can invent one, you can invent an unlimited number. A protolanguage adequate for the needs of hominids two million years ago wouldn't have needed many. Nothing suggests that an insatiable demand for new symbols would have driven the emergence of a phonological combinatorial system.

As Jackendoff is well aware, at least one current framework (Optimality Theory) proposes "a united grammatical framework for syntax and phonology" (Smolensky 1999). Whether or not one buys the theory itself, it seems highly likely that language's two combinatorial systems came in together, perhaps exploiting some single underlying capacity, but more likely with phonology employing mechanisms derived directly or indirectly from syntax. This pushes the third of Jackendoff's stages to a post-protolanguage position.

"Concatenation of symbols" is supposed to constitute another intermediate between call systems and protolanguage. But since "language-trained" apes appear to have concatenated symbols with no explicit training and minimal modeling, why is this stage not implicit in the development of symbols? And why invoke, as a distinct stage, "use of symbol position to convey basic semantic relations"? In every variety of protolanguage I know of, such use is not principle-based but merely a statistical tendency. The real evolution in language was not from unordered symbols to regularly ordered symbols to modern syntax. It was from concatenation in linear strings to concatenation in hierarchical structures (Bickerton 2002). Between these two types there is no intermediate, therefore, not even the possibility of a gradual evolution from one to the other.

Regarding post-protolanguage changes, I have already conceded (Bickerton 2000, sect. 4) that the original two-stage model has to be supplemented by a third stage, the grammaticization of a morphologically bare syntax to enhance parsability. I see no point in arbitrarily dividing this third stage into several sub-stages, as Jackendoff does in his Figure 8.1, especially as Creole languages quickly create both grammatical (albeit unbound) morphology and symbols encoding semantic relations through demotion of regular lexical items. Moreover, each hypothetical stage requires its own selectional history; it will not do merely to suppose that any improvement in a system is automatically selected for.

Whatever its defects, the three-stage model sought to ground itself in known human-evolutionary developments and anchor itself at least provisionally in time. Jackendoff rejects these constraints (explicitly, in the case of time) in the belief that they "make little difference" (p. 236). I'm sorry, they make a lot of difference.

The most striking fact about human evolution is the massive cognitive and behavioral difference between our species and all antecedent species. Moreover, most writers agree that language was strongly contributory to, if not wholly constitutive of, that difference. But if language was evolving gradually over a long period, as Jackendoff's account implies, then why did improvements in language yield no apparent changes in cognition or behavior until the last hundred thousand years?

The gross mismatch between the archaeological record and any gradualist account of language evolution is something that linguists and nonlinguists alike have been studiously avoiding or