

Review of Mark C. Baker "The Atoms of Language" and Ray Jackendoff "Foundations of Language".

Bart de Boer (2003) Book Review of *The Atoms of Language: The Mind's Hidden Rules of Grammar* by Mark C. Baker and *Foundations of Language: Brain, Meaning, Grammar, Evolution* by Ray Jackendoff, *Artificial Life* 9 (1) 89–91

Over the last few years, interest in modeling language has risen within the artificial life community (i.e. Kirby 2002). This work has to do with understanding the evolution of language, the dynamics of language in a population, but also with understanding how language functions in an individual. Unfortunately, this work is little known within the linguistics community, and what little of it is known is often criticized as being too simplistic. This has to do with the fact that linguists do not know much about the problems of modeling, but also with the fact that more than a few modelers know very little of (general) linguistics.

Two recent books that are of interest to researchers (especially those without much formal linguistic training) interested in modeling language are *The Atoms of Language* by Mark C. Baker and *Foundations of Language* by Ray Jackendoff (both of Oxford University Press, 2002). Although both books have rather similar titles, their approaches to the problems of linguistics differ considerably.

Baker's book, *The Atoms of Language: The mind's hidden rules of grammar*, explores one particular theory about language, and tries to make it accessible to an audience of relatively lay readers. It starts with the observation that languages can be astonishingly different, but still translatable into each other. He illustrates this with the story of the Navaho code talkers of the Second World War: the US used speakers of the Navaho language to communicate confidential information about military operations by radio. While the Japanese were unable to understand the messages, or find any way to decode them, the Navaho speakers themselves had no trouble at all translating complex messages to and from English. Baker explains this apparent paradox using the principles and parameters theory. This theory, originally invented by Noam Chomsky (1981) proposes that the grammatical differences between human languages can be explained on the basis of a small number of hierarchically organized discrete principles and parameters. The principles determine the different possible ways in which human grammars can be organized, while the parameters are set by infants learning the language and determine how the grammar will turn out exactly. In the course of the book, Baker identifies a number of potential parameters, using data from about fifty different languages, from Abkhaz to Zapotec. In organizing the material in the book, Baker effectively uses chemistry as a metaphor, and presents the principles and parameters theory as the periodic table of language. In his view, the principles are the "atoms of language" hence the title of the book.

The book does an admirable job of presenting a complex theory of language in very understandable terms and of giving an overview of the complexities of grammar, but it did not manage to convince me (admittedly, not a proponent of the theory to begin with) that principles and parameters are really the key to the understanding of grammar. There seem to be too many exceptions to the grammatical phenomena that, according to the theory, should be rigidly determined by principles and parameters. How can the theory account for such exceptions? Also, Baker's arguments from child language acquisition fall short of being convincing. One can imagine different ways of explaining the phenomena of grammatical development Baker uses to support principles and parameters theory.

However, the book presents a lot of evidence of correlations between aspects of grammar that might not seem correlated at first sight. Also, it shows how the principles and parameters can be used to compactly describe the grammatical differences between languages, and for these reasons it definitely provides food for thought, even for those who are not convinced that principles and parameters are really represented in the brain.

Jackendoff's book, *Foundations of Language: Brain, meaning, grammar and evolution*, is quite different in approach from Baker's book. It contains an enormous amount of material that is the result of a career's long thinking about language and meaning. As such it is a much less easy read than Baker's book, but still it is not just intended for an audience of linguists (Jackendoff says so explicitly in his preface). It provides a broader overview of what language is and the problems that need to be solved for a more complete understanding. It also presents possible theories. Apart from syntax, phonology, semantics and a little bit on pragmatics are also treated. Jackendoff's claim is that all these systems are generative. Traditionally, generative grammarians concentrated on syntax as the only generative component. Jackendoff shows that by taking a broader view on what is generative in language, many phenomena that are considered problematic become less so. He also argues that the lexicon is where all three generative systems are linked: lexical entries are links between grammatical properties, semantic content and phonological form. Jackendoff argues convincingly that this solves a number of problems that plague more traditional generative models of language.

Jackendoff also presents an interesting and well-illustrated possible account of the evolution of language. He presents a number of steps in which language could have evolved from a simple call system and even presents what he calls linguistic fossils to illustrate the different steps. These fossils are expressions that cannot be considered complete sentences, (Hello, Oh boy etc.) utterances such as "shh" or "aargh" but also pidgin-like speech that is used for communication between people that do not speak the same language.

The theories and models are presented in considerable detail and the book abounds with example sentences. Unfortunately, almost all of Jackendoff's examples are of English, weakening his claim that the theory is about language in general. It is not quite clear whether all of Jackendoff's proposed models and mechanisms would be valid for the variety of human language that is, for example, illustrated in Baker's book. This is, in my opinion a weak point of the book, but on the other hand, the depth into which Jackendoff goes into for English could not be reached in a book that has to cover a lot of little known languages.

Both books contribute to a better understanding of the complexity and diversity of human language, and provide an overview of two generative linguistic theories. Baker's book focuses on a specific syntactic theory (principles and parameters) while Jackendoff's book gives a broader overview, and presents semantic theory as well as syntactic theory. As such, both books provide a good introduction into the complexity of language and linguistics, and could serve as a starting point for modelers wishing to learn more about language (although to get a balanced view of what linguistics is about, it would be advisable to read a book from a more descriptive tradition, such as Comrie's classic *Linguistic typology and language universals* (1981) as well).

Both books are relevant to artificial life researchers in a different way as well. They contain both implicit and explicit claims about the implementation of language in the brain. Baker's book does not contain explicit statements about computational models, but the theory of principles and parameters claims that such innate knowledge

simplifies acquisition of language. However, a priori it is not at all clear whether determination of the values of such parameters is any simpler than other methods of grammar induction. This could be tested in a computer model, but has never been done as far as I am aware.

Jackendoff's book is more explicit concerning modeling. The theories presented in the book are said to reflect language's implementation in the brain, refer to processing, representation and interfaces between different modules. It also contains numerous critical references to connectionist models, generally pointing out their insufficiencies and challenging the connectionist researchers to do better. Finally, it also contains a detailed account of possible language evolution, a favorite topic of computer modelers.

Although Jackendoff has a very critical attitude towards connectionist models and it appears as if he does not seem to fully appreciate the difficulties of constructing such models nor the exact aims with which they are generally constructed. But in all fairness it must be said that his criticisms are very much to the point. Especially Jackendoff's critical remarks about the problem of representing two instances of the same object at the same time in a neural network ("the problem of two") and the problem of instantiating variables in a neural network representation are relevant to computer modelers of language. Given that neural networks are universal computers in theory, it should be possible to solve these problems. Although a model large enough for practical language problems would probably be computationally intractable, it would already be very useful to show how these problems can be solved for a small problem. This would address many of the objections linguists have with the use of connectionist models.

On the other hand, Jackendoff himself can be criticized from a computer modeler's point of view. His models have a claim of formality, but nowhere in the book do his models become nearly formal enough to be implemented as a computer program. Ideally, it should be possible to implement a sufficiently formal and complete model of language as a computer program that understands and produces (limited) language. Jackendoff's models fail this test.

But neither the state-of-the-art of computational modeling of language nor the state-of-the-art of linguistics are sufficiently advanced to model language successfully. What is necessary is a more profound collaboration between linguists and modelers (as Jackendoff remarks in his book). Both Baker's and Jackendoff's book contain enough information for modelers to become more fluent in linguistics as well as a number of computational challenges that could be tackled. They are therefore well worth reading for the artificial life audience.

References

- Chomsky, N. (1981) *Lectures on government and binding*, Dordrecht: Foris.
- Comrie, B. (1981) *Language universals and linguistic typology*, Oxford: Blackwell.
- Kirby, S. (2002) Natural Language From Artificial Life, *Artificial Life* **8(2)** 185-215
- Baker, M. C. (2002) *The Atoms of Language*, Oxford: Oxford University press.
ISBN: 0-19-860632-X
- Jackendoff, R. (2002) *Foundations of Language*, Oxford: Oxford University press.
ISBN: 0-19-827012-7