

## BOOK REVIEWS

### THE ACQUISITION OF SYNTACTIC KNOWLEDGE

Robert C. Berwick

Cambridge (MA) and London: The MIT Press, 1985,  
xii+368 pp.  
ISBN 0-262-02226-5, \$27.50

The aspect of human language processing computational linguists are most interested in simulating is the comprehension of individual texts, assuming knowledge of the language that the texts exemplify. There are many reasons why this problem is central to computational linguistics (for one thing, it has obvious commercial applications); but this emphasis is strikingly at odds with the tradition in non-computational theoretical linguistics. There, the process linguists have been chiefly concerned to explain is the child's activity of inferring the structure of a language from examples, assuming knowledge of the range of possible human languages. Theoretical linguists have not been much interested in the parsing problem. (Sometimes, e.g., Fodor, Bever, and Garrett 1974: 368ff., 409, they have suggested that human language comprehension may well not involve a systematic parsing algorithm at all.)

In this book, Berwick brings the concerns of computational and theoretical linguistics together by describing a computationally-implemented algorithm that takes examples of a language as input and learns the language, in the sense that it produces as output a parser for the language (rather than merely a definition of the language in the form of a generative grammar, as in the work of researchers on language learnability such as Wexler and Culicover (1980)).

Concretely, Berwick requires his system to output a parser, having been given a finite and reasonably short, randomly-ordered input sequence of fairly simple example sentences, with no "negative examples" (a child does not encounter ungrammatical sequences with markers to identify them as ungrammatical). The output parser is a version of Marcus's Parsifal (Marcus 1980 – this was chosen, I take it, as a system that claims validity as a psychological model); but, while Berwick retains Marcus's mechanisms of a stack of nodes seeking daughters and a fixed-length buffer filled with nodes seeking mothers, he modifies the system heavily in other respects. For instance, rules are not grouped into "packets", and their relative precedence is determined automatically by their form rather than being stated explicitly as a separate item of information.

The modifications to Marcus's system are made partly in order to simplify the mechanisms by which the parser grows; but, much more important, they reflect thinking among theoretical linguists about universal grammar. It is a cardinal principle of Chomskyan linguistics that

human language acquisition is successful only because much linguistic structure, being innate and common to all human languages, does not need to be learned by the individual. Berwick aims to ensure that features of this kind are built into the fixed rule-interpretation mechanisms of his parser, rather than being included in the rules which are created as the acquisition system runs. The consequence is that the rules of Berwick's parser are formally much simpler and less diverse than Marcus's equivalents. Roughly speaking, a rule in Berwick's system is always an if-then statement in which the "if" side refers to the properties of the top node in the stack and the first node in the buffer, and the "then" side specifies one of four actions: attach the first buffer node as the next daughter of the top stack node; exchange the contents of the first two buffer positions; drop a "trace" into the buffer (a trace is the current Chomskyan device for marking the logical position of a constituent that has been shifted in surface structure); or drop a specified closed-class word, such as *you* or *of*, into the buffer (to deal with cases where logical constituents are deleted in surface structure).

Marcus's parser is deterministic, in that structures once built up by it are never undone. Berwick's language-acquisition system builds up a deterministic parser deterministically. At each step when an input sentence proves unparseable, the system runs through its four parse-rule types to discover and adopt one whose application to the current stack and buffer configuration is compatible with the tight constraints imposed by universal grammar: a rule once adopted may subsequently be generalized, but is never discarded. One of Berwick's claims is that the nature of human language, taken together with a conservative acquisition strategy that always chooses the narrowest form of rule compatible with a triggering datum, is such that the search space of possible parsers relative to a given language has a smoothly-sloping, single-peaked geometry.

Berwick points out that whether an acquisition algorithm exists for a class of languages and whether the languages in the class are parsable are separate questions whose answers do not necessarily coincide. Part of what he claims to have achieved is a demonstration that the set of constraints on language defined by current theories of universal grammar make the class of human languages both a learnable and a parsable class. Like Marcus, Berwick discusses only the parsing of English; but in Berwick's case the parser is determined exclusively by the interaction of English data with a theory of linguistic universals embodied in the parser-acquisition mechanism, and this means that if the theory – which was developed independently of Berwick's work – is correct, then his system should be equally successful at building a parser for any other human language. And Berwick intends his

system as more than a mere existence proof: he offers it as a psycholinguistic theory of how children actually acquire their first language, and he frequently cites observations about child language as tending to confirm his theory.

A book that can make even a *prima facie* plausible claim to have achieved these things must be an important one; hence the length of this review. Just how far Berwick's attempt has succeeded is a question not to be answered quickly. The book is dense, and Berwick is not always as skilled as he might be at helping the reader to disentangle the central skeleton of his exposition from peripheral technical details. An adequate assessment of Berwick's work will need extended consideration by the scholarly community, preferably with further elucidation by Berwick himself.

To set the ball rolling, let me mention some points that worried me on a first reading of the book, though I do so without any assumption that they will ultimately prove fatal to Berwick's case.

My chief worry is that Berwick is quite vague about the computational implementation of his system, and about the degree of success it has attained in practice. The target he sets for his system is to reconstruct the set of rules contained in Marcus's parser, which Berwick describes as numbering "on the order of 100". At one point Berwick states that "by the time the system has processed several hundred sentences, it has acquired approximately 70% of the parsing rules originally hand-written for [Marcus's] parser". Later, he says that "On the order of 70-100 rules . . . are learned from a corpus of several hundred sentences". What is the truth: does the system achieve 70%, or 100%, success with respect to the chosen criterion? If the former, what sort of errors are made? How far does the number of data sentences required tend to vary with the order of data presentation? How naturalistic are the data?

Secondly, although Berwick frequently claims that his theory makes correct predictions about child language, he is again vague about the facts in this domain: ". . . the NP after *by* could be taken, incorrectly, as the Object of the verb. This seems to happen with children"; ". . . children seem to frequently drop Subjects . . . Hyams (1983) has confirmed this . . ." (referring to an unpublished paper which is not discussed further; such remarks are characteristic, and unsatisfying).

This latter point is the more worrying, since it is clear in other respects that Berwick's expertise lies more in the computational field than in the natural language domain. Occasionally he perpetrates linguistic howlers; for instance, he claims that the ungrammaticality of *\*There were a riot on Tuesday* shows that existential *there* governs subject-verb agreement, and he describes the word *assigned* as "trisyllabic". Berwick's weakness in the area of empirical linguistics has the consequence that he is excessively willing to accept every temporary theoretical proposal of the M.I.T. school of linguists as gospel, fail-

ing to distinguish long-term, core principles from trial balloons which someone floated last year and which will probably be abandoned next year. In Part II of his book, which discusses the theoretical implications of the acquisition model as opposed to its internal workings, Berwick commits himself to a number of linguistic beliefs that seem indefensible. He spends some time discussing a constraint on natural-language semantics (attributed to F. Sommers and F. Keil) according to which graphs representing the relationship of predictability between vocabulary items rarely or never contain *M*-shaped subgraphs. This seems quite wrong (a *kitten* and a *baby* may both be *stillborn*, a *baby* and an *engineer* may both be *British citizens*), and it is not clarified by the diagram Berwick uses to illustrate it (p. 270). He devotes many pages to arguing that his theory explains an alleged phonological constraint, quoted from an unpublished doctoral dissertation by M. Kean, which (as Berwick describes it) appears to forbid the occurrence of a language having stop consonants at three or more places of articulation but a fricative at only one. In reality, systems with /p t k/ and /s/ but no /f/, /x/, etc. are rather common.

These aspects of Berwick's book seem regrettable, and unnecessary. The book would have been a significant contribution if most of Part II had been omitted. I wish it had been, and that the computer implementation had been discussed more fully; but I hope these points will not cause the valuable parts of the book to be neglected.

*Geoffrey Sampson*

Department of Linguistics and Phonetics  
University of Leeds  
Leeds LS2 9JT England

## References

- Fodor, J.A.; Bever, T.G.; and Garrett, M.F. 1974 *The Psychology of Language*. McGraw-Hill.  
 Marcus, Mitchell P. 1980 *A Theory of Syntactic Recognition for Natural Language*. MIT Press.  
 Wexler, Kenneth and Culicover, Peter W. 1980 *Formal Principles of Language Acquisition*. MIT Press.

---

## STRUCTURES AND PROCEDURES OF IMPLICIT KNOWLEDGE (Advances in discourse processes 17)

Arthur C. Graesser and Leslie F. Clark,

Norwood, NJ: Ablex Publishing Corp, September 1985,  
viii+326 pp.  
ISBN 0-89391-192-5, \$42.50 (cloth); ISBN  
0-89391-362-6, \$24.50 (pbk)

How does one understand a narrative? The current scientific theories reduce comprehension to the generation of a correct sequence of inferences from some knowledge structures. Inferences are obviously created when individuals comprehend text, but there is widespread disagreement about what inferences are generated, when