and English. His extension is to my mind ad hoc and although it works for his examples, might not adapt itself to more complex problems like negation in French.

Karen Jensen ("Binary rules and nonbinary trees: Breaking down the concept of phrase structure") is concerned with the passage from binary rules (and trees) that capture significant generalizations about natural languages to list structures that are more satisfactory for further processing. Her solution can handle discontinuous constituents and is suitable for treating languages with free word order.

In "The notion of 'rule of grammar' reconsidered" Michael Kac defends the notion that "grammatical analysis requires that we have a way of formally representing the variety of distinct etiological properties that can be manifested by ungrammatical strings, this diversity corresponding to the variety of distinct rules of grammar" (p. 137). He argues that getting the standard linguistic theories (various versions of TG, GPSG, etc.) to serve the purpose of etiological analysis is "problematical". His "fundamental principles" (pp. 120, 122) appear to require that a grammar supply a structure not only for elements of the language L that it generates but also for the elements in the complement of L. In his formal development, however, he defines an object (Definition 9) in terms of itself and this circularity would appear to render the result ill-defined. Since his main argument depends on this definition, I stopped reading. It is a good practice to buttress complicated definitions with examples both for the good of the writer as well as that of the reader.

There are three papers on tree-adjoining grammars: an introduction by Joshi, "Unbounded dependencies and subjacency in a tree adjoining grammar" by A.S. Kroch, and "On the progression from context-free to tree adjoining languages" by Joshi et al. This presents an easy access to a useful collection of results concerning a rather pregnant linguistic model.

Finally, three of the papers are concerned with semantics proper. G.N. Carlson's "Exceptions to generic generalizations" deals with the construction of a formal semantics using a sort of default mechanism in order to interpret statements like "Dogs bark" when clearly barkless dogs exist. Davis and Papcun in "The structure underlying a semantic domain" provide a rather metaphorical model *vr* (volumetric representation) "to formalize lexical knowledge in a practical way". They investigate various models—semantic networks, multi-dimensional scaling, and clustering—before settling on their own spatial (and somewhat intensional) model of a semantic domain. The third paper, R.H. Thomason's "Remarks on linguistic semantics", is an expository article concerned with the interface between linguistics and philosophy, dealing with the literature of such topics as tense and aspect, propositional attitudes, and vagueness.

Barron Brainerd is Professor of Mathematics and Linguistics at the University of Toronto. He is the author of *Weighing Evidence in Language and Literature: A Statistical Approach* and *Introduction to the Mathematics of Language Study*. Brainerd's address is: Department of Mathematics, University of Toronto, Toronto, Ontario, Canada M5S 1A1.

The Vastness of Natural Languages

D. Terence Langendoen and Paul M. Postal
(Graduate Center, City University of New York and IBM Thomas J. Watson Research Center, Yorktown Heights, NY)

Oxford, England; Basil Blackwell, 1984, ix + 189 pp. ISBN 0-631-13461-1 (hb); ISBN 0-631-14756-X, £8.95 (sb)

Reviewed by
James V. Rauff
Millikin University

This book is an extended argument in support of the theses that natural languages are transfinitely unbounded collections, that sentences are not limited in length (number of words) by any cardinal number, finite or transfinite, and that no constructive grammar can be an adequate grammar for any natural language.

Chapter 1 is an introduction to those aspects of set theory needed to develop the main points of the book. Specifically, the notion of a class arising from Cantor's and Russell's paradoxes and the Cantor power set are introduced.

Chapter 2 sets forth what the authors call the "received position about natural languages" (hereafter NLs). The received position is that NL sentences are finite in length. Length is defined in terms of number of words, although the authors argue later that we could just as well count phonemes as words without seriously affecting their arguments. NLs as collections of finite-length sentences are therefore countably infinite (or denumerable). Finally, related to the finiteness of sentences is the "received position" that grammars for NLs are constructive.

Chapter 3 argues that there is "no motivation for imposing size laws on NL sentences" (p. 44). Invoking Occam's Razor, the authors claim that size laws are extra-linguistic restrictions not needed for grammatical description and therefore unjustified.

Chapter 4 presents the main theoretical points of the book. Taking as axiomatic for NLs a property of coordination that allows for unrestricted coordinate compounding of sentences, the authors present the **NL Vastness Theorem,** which asserts that NLs are not sets, but rather classes with no fixed cardinality. The argument can be illustrated with their example (pp. 55–57):

1. Let L be the NL English.
2. The set S_0 is contained in L, where
   S_0 = {*Babar is happy*; *I know that Babar is happy*; *I know that I know that Babar is happy*; . . .}

3. S_1 may be constructed as follows.
   a. Form the set of all subsets of S_0, $P(S_0)$.
   b. For each element B in $P(S_0)$, form the sentence that is the coordinate conjunction of all the sentences in B.
   c. Let S_1 be the collection of all sentences formed in (3b).
   S_1 = {*Babar is happy*; *I know that Babar is happy*; *I know that I know that Babar is happy*; . . . ; *Babar is happy and I know that Babar is happy*; *Babar is happy and I know that I know that Babar is happy*; . . . ; *Babar is happy, I know that Babar is happy, and I know that I know that Babar is happy*; . . .}
4. S_0 is denumerable, but S_1, which is equinumerous with $P(S_0)$ is not denumerable (by Cantor's Theorem).
5. S_2, S_3, etc., can be constructed analogously. Each successive S has a greater transfinite cardinality than the one preceding it.
6. All of the S collections are contained within L.
7. L has no fixed cardinality.

Although this particular **construction** shows that the cardinality of L is not fixed by any weakly inaccessible cardinal in the sequence $\aleph$ (see Monk 1969), the authors later imply that their result holds for strongly inaccessible cardinals as well. A notable corollary to this example is that sentences in L have no upper length bound. Indeed, at least one sentence in S_1 has length c, the power of the continuum.

Chapter 5 presents some implications of the NL Vastness Theorem. The implication of most interest to computational linguists is probably the **NL Non-Constructivity Theorem** (p. 72), which states that no NL has any constructive grammar. This theorem, Langendoen and Postal assert, immediately invalidates any generative, proof-theoretic, or Turing machine grammar for natural languages. The authors list 27 invalidated theories (including LFG, GPSG, GB, systemic grammar, and tagmemics). All that remain uninvalidated are Johnson and Postal's (1980) arc pair grammar (which is mentioned) and perhaps Foley and van Valin's (1984) role and reference grammar (which is not mentioned).

Chapters 6 and 7 assert that transfinite sentences do not contradict any linguistic principles of sentencehood, nor is the concept of transfinite sentences invalidated by any existing linguistic ontology, conceptualism, or platonism. Finally, the authors assert that no linguistic principle relevant for the characterization of finite sentences fails to be relevant for transfinite ones.

The vastness of natural languages is a fascinating monograph that takes the formalist approach to human languages to its absurd, yet logical, conclusions. The arguments supporting the NL Vastness Theorem are logically valid, but there continues to be something not quite right about statements like "consider then a denumerably long transfinite German sentence with infinitely many underlying voiced word-final stops" (p. 76). This conflict between what can be logically deduced and what actually is goes to the heart of this book and its relevance for formal linguists and computational linguists.

Langendoen and Postal offer here a kind of reverse parallel reaction of the kind Brouwer and the intuitionists offered against formalism in the foundations of mathematics (see Hatcher 1982). The intuitionists attacked the notion that a contradiction resulting from the assumption of the nonexistence of a mathematical object entailed a proof of its existence. Langendoen and Postal heartily embrace this principle and apply it to natural languages. In this respect, this book offers a much-needed serious discussion on the foundations of formal linguistics. It shows that within the formalist paradigm, theory builders and natural language processing program writers are logically at odds with each other. Computational linguists must, by definition, be constructivists. On the other hand, theory builders, as is shown in this book, cannot be.

For the computational linguist, one important criticism of this book lies in the relationship between natural languages and Langendoen and Postal's NLs. It is not clear that the set-theoretic NLs are indeed natural languages. Perhaps a human language, like English, is no more than a restriction of some sort of the NL-English. Indeed, NL-English may have sentences with length the cardinality of the continuum, but the English that is used by humans or modeled by computer programs need not. Even if all constructive grammars of English are invalid as grammars for NL-English, these grammars may be quite adequate for expert systems, speech processing, etc.

The vastness of natural languages is a well-argued and thought-provoking book. It should be of interest to anyone interested in the foundations of linguistic theory, although its effects on computational linguistic practice may be minimal.

References

Foley, W.A. and van Valin, R.D. Jr. 1984 *Functional Syntax and Universal Grammar*. Cambridge University Press, Cambridge, England.

Hatcher, W.S. 1982 *The Logical Foundations of Mathematics*. Pergamon Press, Oxford, England.

Johnson, D. and Postal, P. 1980 *Arc Pair Grammar*. Princeton University Press, Princeton, NJ.

Monk, D. 1969 *Introduction to Set Theory*. McGraw-Hill, New York.

James V. Rauff is an associate professor of mathematics at Millikin University. He holds a master's degree in mathematical logic and a doctorate in computational linguistics. His recent research has been in two-level grammars and machine translation. Rauff's address is: Department of Mathematics, Milliken University, Decatur, IL 62522.