

ELEMENTS OF MATHEMATICAL LINGUISTICS

(Janus Linguarum, series maior, 110)

A.V. Gladkij and I.A. Mel'čuk

edited by John Lehrberger

Mouton Publishers

Berlin, New York, Amsterdam, 1983

Pp. xxi, 247. ISBN 90-279-3118-6

1983, P. vii, 90.

Most of this book (the first 148 pages) is an English version of an introduction to mathematical linguistics originally published in Russian in 1969. This material has been expanded into a more standard-sized book for the English-speaking market by tacking on three articles on special topics in mathematical linguistics published elsewhere by one or both co-authors; but each of these articles is already available in English translation, and the book stands or falls by the usefulness of its first part.

When Gladkij and Mel'čuk wrote their book in the late 1960s, mathematical linguistics was little known, even in the USA where most of it had been created not long before. I guess that it was a fairly remarkable achievement on their part to discover and master this exotic discipline (and, in Gladkij's case, to make a number of original contributions to it), and then to make it known to their fellow-countrymen.

However, whether it makes good sense to bring out an English version of the book in the 1980s is a different matter. I must say that I feel the answer is no, for many reasons.

In the first place, mathematical linguistics has moved on since the 1960s. Gladkij and Mel'čuk's book deals almost exclusively with the hierarchy of types of language, defined in terms of types of grammar (unrestricted rewrite systems, context-sensitive grammars, context-free grammars, one-sided linear grammars), with a small amount on the relationship with types of automata, on decidability theorems, and a few other matters. The topic of parsing – and mathematical formalisms with

special relevance for automatic language processing, such as the ATN – are not mentioned. In 1969, one would not have expected them to be; in the 1980s, these are surely indispensable components even in an introduction to "mathematical" (as opposed to "computational") linguistics. Gladkij and Mel'čuk are quite explicit about the fact that they have made no attempt to bring the book up to date: "The manuscript of the Russian version of this book was completed in 1967 and we are not in a position to revise it now. . . . Even a mere list of references would be out of the question."

Secondly, the book relies heavily on Russian-language examples which are lost on an English reader. Thus, early chapters largely revolve round a largish formal grammar produced by the authors in order to generate the complex range of Russian participles. It must have been a significant virtue of the original book that it demonstrated how the concepts of mathematical linguistics could be made to achieve a novel task relating specifically to the readers' own language, but for a British or American student the result is that ideas many people find difficult at the best of times are rendered wholly opaque.

And, finally, Gladkij and Mel'čuk just are not very good at writing for an unknowledgeable audience. They make their formal rules notationally much more complex and exotic-looking than they need be – to a fellow mathematician, a trivial matter, but for an unconfident student very unfortunate. On page 25 they use the term *monoid* without, I think, ever explaining what it means (there is no index, so it is hard to check); on page 81 they use a technical term of their own which is first defined on page 122.

The fact is that there are now enough English-language books that do the same job as this book and do it much better.

Geoffrey Sampson

Department of Linguistics and Phonetics

University of Leeds

LEEDS LS2-9JT

**AUTOMATIC SEMANTIC INTERPRETATION:
A COMPUTER MODEL OF UNDERSTANDING
NATURAL LANGUAGE**

Jan van Bakel

Foris Publications

Dordrecht, Holland, 1984

Pp. x, 176, \$14.40, Dfl. 36.00

ISBN 90-6765-039-0

"Semantic interpretation" is an expression that seems to mean something different to everyone who uses it. What it means to van Bakel is translation from a Dutch surface form into a case-like structure, which is expressed in a language called SELANCA. The interpretation retains the

content words but not the structure of the input; that is, it is not an expression in an interlingua or knowledge representation that is independent of the source language.

Van Bakel takes view of both language and linguistics counter to those currently prevailing in North America (I don't know about Europe or elsewhere). First, he rejects any connection between theoretical linguistics and computational linguistics, seeing the former as no longer having anything to say to the latter, which may now proceed solely as an application-oriented enterprise. Second, in contrast to the situation semantics view à la Barwise and Perry (1983), van Bakel sees reality as being structure or constrained by language rather than the other way round. "The reality which is related to [a

word such as] *fat* would not exist as such if it were not created by language. There is no ground to postulate an autonomous reality that is referred to 'by natural language" (pp. 35-36).

Either in consequence of these views, or despite them – I can't decide which – van Bakel explicitly rejects semantic interpretation that "manipulates reality or . . . a model of reality" (p. 33), such a database update and retrieval:

I do not want a certain part of the world to be introduced within the research situation, as is the case in question answering systems, . . . because I want to stick to linguistic theory.

. . . Linguistics is passed when reality is reached. Connecting language and reality will not be the task of linguistics (p. 32).

My study should not be considered as a contribution to meaning theory or semantics. What the reader should expect is an exposition of a certain view on meaning, together with a view on semantic interpretation, and an explanation [of] how this is performed in the model described (p. 32).

Notwithstanding these remarks, van Bakel makes strong claims for his model: "The . . . system is claimed to be a model of a native speaker's competence of understanding natural language sentences" (p. 1). However, there is little discussion of the adequacy of SELANCA language as a representation, even on the author's own terms; rather the emphasis is solely on the translation process.

The model itself consists of a context-free parser, AMAZON, which produces a syntactic structure that is then converted by the semantic component, CASUS into a SELANCA expression. There is no feedback from semantic to syntax; the user is asked to resolve syntactically ambiguous words. If more than one parse is possible, AMAZON will give all structures; and CASUS may find more than one interpretation for a given structure. There is no attempt to solve the ambiguity.

Van Bakel goes into considerable detail about the operation of CASUS. The interpreter first performs transformations on the parse tree for detopicalization, eliminating passives, and so forth. It then assigns a case structure to the sentence, using morphological and simple lexical features. The interpretation of sentential complements is given particular attention.

The examples are mostly in Dutch, with literal English glosses. A knowledge of Dutch is not necessary for reading the book, but it would be helpful, especially for the detailed examples of the appendices. There are a few places where more examples would have helped; for instance, the SELANCA language is introduced with no examples at all. In the appendices, SELANCA expressions are rendered impenetrable by their not being pretty-printed.

I have great admiration for authors such as van Bakel who can write so fluently in a language not their own. However, the book would have been improved if copy-edited by a better speaker of English. The reader is misled a little too often by misplaced commas, by the consistent use of *it's* for *its*, and by the occasional wrong word. Nevertheless, the author's pleasantly informal style comes through, even if the argument is a bit hard to follow at times.

My qualms about this book derive mostly from the author's unorthodox views, which he presents strongly but doesn't always stick to himself. He believes he is doing computational linguistics, strictly application-oriented; yet he seems to dislike applications such as database interfaces and presents his work as a linguistically interesting general mode of understanding, and doesn't want to constrain his work to subsets useful for interfaces (p. 3). He claims the work contributes nothing to the theory of meaning, and yet calls it semantic interpretation rather than, say, case analysis. And he presents it as a model of "human understanding" without any discussion of what human understanding is or why the model models it.

What's good about the work is that it is a fine detailed description of a parser and case analyzer, a system that may indeed be the prototype of a useful natural language interface.

Graeme Hirst

Department of Computer Science
University of Toronto
Toronto, Canada M5S 1A4

References

- Barwise, Jon and Perry, John R. 1983 *Situations and Attitudes*. The MIT Press / Bradford Books, Cambridge, Massachusetts.