

English. This is particularly so in the early part of the book. When making a proposal, they often examine several possibilities which are all plausible in their framework. This should be welcomed by readers with only limited exposure to formal semantics.

The authors address methodological issues at several points, since their main concern is the presentation of a fragment. Nevertheless, it would have been useful for the authors to situate their work more explicitly by indicating what a boolean emphasis has added to semantics, what areas are likely to benefit from a similar approach, and what problems will require new methods. This book shows that the algebraic approach offers much of interest to semantics.

*Lawrence S. Moss*  
Department of Mathematics  
University of Michigan  
Ann Arbor, Michigan 48109

#### GENESIS: AN AUTHORSHIP STUDY IN COMPUTER-ASSISTED STATISTICAL LINGUISTICS

**Y.T. Radday and H. Shore**  
with **D. Wickman, M.A. Pollatschek, Ch. Rabin, and**  
**Sh. Talmon**

Rome: Biblical Institute Press, 1985, xx+235 pp.

This book studies a topic in Bible scholarship by means of computer-assisted statistical methods. The book has seven parts:

- I Introductory, by Radday,
- II On statistics in general and in Genesis, by Wickman,
- III Statistical analysis of formal criteria, by Shore and Radday,
- IV Vocabulary richness and concentration, by Pollatschek and Radday,
- V. An interim postscript, by Radday and Shore,
- VI Linguistic aspects, by Rabin,
- VII A Bible scholar's evaluation, by Talmon.

The book combines the work of authors from different fields of research: Radday is a Bible scholar, and as a matter of fact, was the first to introduce statistical computational methods into Bible study in Israel; Pollatschek is specializing now in operations research and computer techniques. Both are at the Technion. Shore carries titles in industrial engineering and operations research, as well as in philosophy and psychology. Wickman teaches mathematics and statistics at the Technische Hochschule in Aachen, FRG. Rabin is from the field of classical and modern Hebrew and Arabic linguistics, and Talmon is Magness Professor of Bible Studies. Both these latter scholars are from the Hebrew University of Jerusalem.

The goal of this study was to examine the authorship of the Book of Genesis. Textual and exegetical diffi-

culties in this book roused theories, such as Wellhausen's, that the Book had not been written originally by one hand, or that a later editor edited material written by several (at least three) previous writers. This "documentary theory" is investigated in the study in three phases:

- Phase I general statistics;
- Phase II statistics of linguistic data;
- Phase III vocabulary statistics.

The text of Genesis was divided into three parts:

- (1) Sort of Text:
  - N (Narrative);
  - H (Human speaker);
  - D (Divine speech);
- (2) Documentary Source (following Wellhausen's theory):
  - J (text pieces using the letters JHWH for God),
  - E (text parts using the word Elohim for God),
  - and
  - P (representing a priestly writer);
- (3) Division I, II, III, according to story-type, namely the first cycle of stories of the creation, the flood etc., the heroic stories of the Fathers, and the cycle of stories about Joseph. The creation story of the first chapter and Jacob's blessing were excluded since they were linguistically too deviant.

Further slicing of each of the above classes yielded text sequences of about 200 words each, which were convenient for statistical analysis. Each sample was statistically examined for 54 linguistic items in the areas of word length (2 to 10 characters); certain nominal and verbal morphology elements; syntactical elements; and frequency of inter-word transfers, such as noun/noun, noun/verb, noun/pronoun, and noun/stop. These features are unique for any writer, and cannot be consciously manipulated. They therefore reveal the writer's individual style and may corroborate assumptions concerning the text's author. Statistical results show minute differences per item; but when all details are collected, consistent facts of certain linguistic features emerge and yield a complex picture of the linguistic structure of the text.

In Phase I the univariate analysis of variance revealed that J and E were indistinguishable from one another, while P was strikingly unlike either. Also, NP was heterogeneous, while NJ and NE were not. The multivariate analysis of variance demonstrated the same pattern of differences among the documents. P appeared to be of an independent source, while J and E bore very close resemblances.

In the analysis of the sorts of discourse, N is completely unlike H and D in any document. Seventeen out of 39 variables were found to be powerful discriminants between N and H+D. All the results of these analyses mean that the Narrator behaves linguistically in a significantly different fashion from the speakers.

From the aspect of divisions, it is shown that Division I is quite unlike Divisions II and III. In Division I were included most of J and P, and in Division III most of E. E appears for the first time in Division II, and from there it has approximately the same share as J. Division II is gradually and increasingly individualized, and fully human true-life portraits appear in Division III, while Division I has a quasi-mythical nature.

In Phase II, the aim was to let the features arrange themselves into groups of themselves, as it were, without any preconceptions. Five analysis methods were applied, and cluster analysis yielded a most interesting picture of interrelations among the text samples (figures 3.11, 3.12, pp. 136-7). The first cluster comprised samples of Text P. The second cluster comprised most of the N samples outside P. The third cluster encompassed all the H and D samples, regardless of their E or J origin. The authors write explicitly (p. 186): "This adds force . . . to the impression that the Jahwist and the Elohist were each other's alter ego". The other analyses supported this finding.

Lexical richness and lexical concentration were separately examined, yielding results in three different ranges of vocabulary for N,H,D (Fig. 4.3) and J,E,P (Fig. 4.4). Here, too, P is unlike the other groups. After the statistical analysis the authors consider it justified that N,H,D should be distinct. It is doubtful, to their mind, whether J,E,P should be ascribed to three different sources. They summarize this issue as follows: "While solely on these grounds this [documentary] hypothesis cannot be rejected out of hand, such serious doubts regarding its validity have arisen that neither can it any longer be accepted as unreservedly as it has been hitherto." (p. 214)

Rabin in his contribution is not surprised that the sources of discourse have been found so distinctive. Indeed, modern linguistics tends to take into account various extra-linguistic (psycholinguistic and sociolinguistic) factors that yield distinct linguistic patterns in dialogue, narration, etc. Narrative is known to differ from direct speech in various linguistic elements, such as use of pronouns, verb tenses, and demonstratives. It is, then, little wonder that Deity should use an elevated style, full of rhetorical elements on the one hand, and complete and complex sentences on the other, as found in Genesis. He suggests examining the Book according to discourse analysis techniques and the comparison of results.

Talmon warns the reader against the spell of modern technological aids, since another computer-aided study of a part of Genesis by a French scholar has yielded different results. He notes that N predominates all over the text of Genesis, with 53% of Division III, 56% of Division II, and 74% of Division I. H increases from 5% in Division I to 34% in Division II and 47% in Division III. D, however, decreases from 21% in Division I to 10% in Division II and complete absence in Division III. This

internal consistency leads him to suggest that N may have used pre-existing Canaanite or generally known Mesopotamian material. Talmon stresses that the kind of text analysis as applied in this book is really a literary analysis, which fits well with modern trends of Bible scholarship.

Belonging to the linguists' part of the expected reading public (which according to the writers includes "Bible scholars, linguists, statisticians, and probably computer people"), I find this book interesting to study. The starting point of the documentary hypothesis is concisely and clearly summarized. The statistical methods of each phase are clearly described, and are illustrated by tables, figures, and appendices. The findings appear at the end of each statistical analysis, with a brief summary of the results (without the numbers). It may be noted that there is no list of references, and references to the literature appear as footnotes. Appendices are said to be four in number, but only three exist, although "appendices 3 and 4" comprise various sub-sections. There is also a printing mismatch of the number of counted words between the text (p. 192) and table 4.1 (p. 193) and exhibit 4.1 (p. 129).

The results of this study make us hope this is not the last word on the subject (see "Suggestions for further investigation and two examples" pp. 187-189, and "Interim Postscript" pp. 216-217). Subsequent studies, perhaps with more details of the linguistic features proper, are to be awaited.

*Judith Rosenhouse*

Department of General Studies  
The Technion, Haifa, Israel

---

#### GETTING COMPUTERS TO TALK LIKE YOU AND ME: DISCOURSE CONTEXT, FOCUS, AND SEMANTICS (AN ATN MODEL)

**Rachel Reichman**

Cambridge, MA: The MIT Press, 1985, xiii+221 pp.  
ISBN 0-262-18118-9; \$20.00

In this optimistically-titled book, Reichman presents a theory of discourse processing that is intended to be the basis for a computational module. The claim is that this would give a full-scale natural language processing system the capacity to generate and interpret extended coherent discourse, rather than just short utterances. The module would be one of the many that make up such a system, although Reichman does not consider what other components would be required, and how they would interact with the discourse module.

The module would keep a dynamically-updated record of the discourse it is engaged in, in such a way as to help it to produce conversational moves that are relevant to the discourse and contain correct use of pronominalization and other discourse phenomena. Reichman provides