PART TWO

METHODOLOGY OF THE
COMPUTER-ASSISTED LINGUISTIC
ANALYSIS
CHAPTER SEVEN

THE CALAP MODEL OF TEXTUAL ANALYSIS

7.1 INTRODUCTION

In the CALAP project, of which the present monograph is a product, we have developed a model of linguistic analysis and text interpretation, in which both insights from linguistics, especially computer linguistics, and text-critical and text-historical considerations are taken into account. The present chapter will be concerned with the characteristics of this model and its underlying assumptions concerning languages and texts. We will also discuss the implications of this model for the analysis of texts and translations.

The way in which a text is approached in a computer-assisted analysis differs considerably from that in the traditional philological analysis. From a computer-linguistic perspective a text is a one-dimensional sequence of characters. Behind this string of characters several layers or dimensions of information can be added. It is also possible to mark relationships between non-sequential elements, or to take into account several witnesses of one text as parallel sequences of characters. However, the understanding of the text as a one-dimensional entity remains radically different from the philological understanding of the text as an abstraction, a scholar construct on the basis of the extant manuscripts and quotations, which are the result of a long and complicated transmission history.

---

1 CALAP stands for Computer-Assisted Linguistic Analysis of the Peshitta. For more details about this project see www.leidemuniv.nl/gg/calap. For its background and methodology see Van Keulen–Van Peursen, Computer Linguistics and Textual History. See also preface to the present study.

2 Cf. Kroeze, 'Multidimensional Linguistic Database'.

3 The database model used in CALAP is a further development of C.-J. Doedens' Monads dot Feature (MDF) model, described in his Text Databases; see also Petersen, 'Emdros'.
