In Search for Patterns. Analyzing Syntactic Diversity in the Hebrew Bible
Wido van Peursen

The linguistic diversity of the Hebrew Bible or Old Testament has sparked much interest over the last decades, leading to a broad range of explanations such as differences in time of origin, genre, dialect, the influence of neighbouring languages, individual styles of the authors, or transmission history. Fierce debates have led to a flow of publications. The significance of many contributions to the debate is tempered by limitations such as a focus on only one or a few of books, methodological presuppositions (e.g. the research question as to whether a certain book of the Bible is linguistically late presupposes that there is something such as ëLate Biblical Hebrewí, Q.E.D.) and a focus on lexical items set phrases or morphosyntactic phenomena.

In 2013 Dr Janet Dyk and Prof. Wido van Peursen of the Eep Talstra Centre for Bible and Computer (ETCBC), VU University Amsterdam, started an NWO project that aims to compensate for these limitations by (1) studying the complete Hebrew Bible rather than separate books, (2) focusing on syntax, and (3) analyzing the distribution of phenomena according to a wide range of parameters before interpreting them in terms of language development, genre etc.

Ad 1. Covering the complete corpus (ca. 420.000 words) is now feasible thanks to the linguistic database of the ETCBC. Some extra-biblical texts (inscriptions, Dead Sea Scrolls, Rabbinic texts) will be included in the corpus as points of comparison. Since this database contains information at various linguistic levels (words, phrases, clauses, sentences, text), it allows for the syntactic analysis envisioned.

Ad 2. Syntax is extremely important in the discussion of language variation and language development in the Hebrew Bible. While words can be borrowed and phrases repeated, differences between syntactic structures point to options within the underlying language system and are therefore less conscious and less apt to be manipulated.

Ad 3. For the analysis of the distribution of the syntactic phenomena we can make use of markings in the database as to syntactic context (independent versus dependent clauses), literary context (narrated text versus direct speech), distribution over corpora, chapters, etc. Such differentiations will bring to light possible factors influencing the observed variation: inscriptions could provide indications of geographic dialects; post-Biblical texts (Dead Sea Scrolls and Rabbinic texts) could give data indicating diachronic development; variation between prose and poetry texts could indicate a difference between genres; differences between portions of the canon, e.g., between the Torah and the rest of the Hebrew Bible, could point to a more conservative versus a less conservative attitude in scribal transmission.

In the paper we will argue that that the use of a linguistic database causes shifts in analytical procedures. It will present the Hebrew Bible project as a case study to show how in fields of linguistic and philological research that have a long venerable tradition of scholarship digital approaches can complement traditional research in this field of study. It will also address the question as to how to deal with their congruous and contradicting tendencies when various linguistic layers are examined.