Finally, one last glimpse into one more aspect of Genesis investigated. The reader may have wondered why the number of variables was given as 56 in one place and 54 in another. The answer is that two criteria were treated separately and did not participate in all the above. They deal with the lexical side of the controversy.

Much has been made of vocabulary by unitarians and Documentarians alike. The recurrence of *qedoš yiśra'el* in both halves of Isaiah is said to be a sign of that book's unity[7] while critics saw in Genesis the alternate use of *'amah* and *šipḥah* reason enough to ascribe the respective text blocks to different writers. Yet either argument is most tenuous. The first may be a case of quotation and the second depend on circumstances, just as in English the same husband may call a certain woman his wife, his spouse, his better part or else. Any text property lying under the conscious control of a writer must be disqualified in an authorship study. What, on the other hand, is indicative of a writer are two quantitative values of his vocabulary: its richness (VR) and its concentration (VC). The former refers to the percentage of *hapax legomena* (or perhaps also *dislegomena*) in his opus, the latter to the percentage of the text size taken up by, say, the first 10 (or perhaps 20 etc.) most frequent words. The obstacles on the way to using either as an author-specifying criterion are many. Why should measuring VR stop at *dislegomena* and not include *trislegomena* too? Why should measuring VC not extend to the first 30 most frequent words? Are not both functions of the text size? And why should the main part of the frequency list, that is the one lying between the most frequent and the most rare words be neglected? Various ways out of this quandary have been suggested, none, though, is satisfactory.

On these grounds, we employed an altogether new theoretical statistical distribution proposed by Sichel which takes into account the frequency list as a whole and makes it independent of text lenght.[8] Accordingly, frequency lists were drawn up by the computer of lemmata for each of the six corpuses J, E, P, N, H, and D separately, however, after each was arbitrarily divided into two equal halves. In this manner, twelve points obtained on a grid where VC may be read on one and VR on the other axis. What emerged for the SDS-dimension may be viewed in Fig. 6.

The Figure is selfexplanatory: the three areas are very remote from each other. If space permitted it, the grid for the DOC-dimension should also be displayed for comparison. Since this is impossible, we must be believed: there, the areas of J and E are close whereas that of P stands out by its high VR. This feature is rather astonishing for of all the three presumed writers it is exactly the Priestly one who is said to be repetitive!

---

[7] This argument for the unity of Isaih is used by A. Kaminka, Meḥqarim ba-Miqra uva-Talmud uva-Sifrut ha-Rabbanit (Tel-Aviv: Devir, 1935).

[8] For this distribution, see H.S. Sichel, »On a Distribution Law of Word Frequencies«, Journal of the American Statistical Association 70 (1975), 542–7.

GENESIS, BOOK OF

ANALYSIS OF THE BOOK OF GENESIS