

# The homogeneity of genesis

by

Y. T. RADDAY

*Israël Institute of Technology*

795

This paper is a highly condensed report on no more than one part of a comprehensive study undertaken and recently accomplished by a group of four : H. Shore, Tel-Aviv University, M.A. Pollatschek, Technion, D. Wickmann, Technische Hochschule, Aachen, and myself.

The problem posed is one of disputed authorship or homogeneity which in itself is hardly different from similar ones already treated by others except in three regards :

- a) We operated not with one variable as did others, e.g. with the incidence of the Greek particle *καί* or with the number of syllables, but with more than 50.
- b) Sophisticated techniques were applied, probably for the first time in linguistic analyses - see below.
- c) The most prominent difference lies, however, in the text itself, i.e. Genesis. Who after all cares whether, for instance, Hippocrates did or did not write a certain treatise ? Our text, on the other hand, is held in reverence by hundreds of millions, Christians and Jews alike, concerns their deepest convictions and guides their daily lives. To examine whether it is ancient and of one piece, as Jewish and Christian traditions have it, or, as suggested by the impressive acumen of Higher Biblical Criticism, a compilation by a late editor from various pre-Biblical sources is therefore more than just another academic exercise.

These critics discovered in Genesis stories that contradicted each others and others that repeated the same theme two or even three times. The most blatant inconsistency is the use of two different names for Deity : the Tetragrammaton JHWH and Elohim. One instance of a repetition is that Sarah was taken by force to Pharaoh's palace in ch.12 and again to Abimelech's palace eight chapters later.

On these grounds, the Documentary Hypothesis, generally named after the German Protestant theologian Julius Wellhausen (1844-1918), postulates that a late 'Redaktor' who can not have lived earlier than 200 years after Moses composed Genesis by incorporating in it ancient pre-Biblical Documents. As these were of course never found, they are called, for lack of anything better, in accordance with the varying names of Deity, J and E and their authors the Jahwist and the Elohist, both, it is generally acknowledged, highly gifted storytellers. However, some parts of Genesis are not dramatic tales, but long rosters of names and clans, long-winded treaties and the like, and thus out of the Jahwist's or the Elohist's characters. Hence, a third writer was surmised who must have been an official, and since the officials were priests, Wellhausen called him The Priestly Writer and his contributions in the book are marked by P.

The tripartition of Genesis into J, E, P is today widely accepted in Biblical scholarship and taught in all universities on earth. Opponents are said to be religiously prejudiced and their reasoning is disregarded. In fact, the two camps are not on speaking terms. It must be admitted that both sides are tainted by a-priori stances and rely heavily on impressionistic arguments. Among these, one is style. Style, though, whatever it may mean, is closely linked with content and therefore precarious ground in authorship enquiries.

It is at this point that our investigation started. Its aim was to assess the validity or not of the Documentary Hypothesis by means of statistical linguistics, based exclusively on formal, countable and content-free criteria not lying under the conscious control of the writer. They include word length in terms of phonemes, the use of the definite article (optional in Hebrew) and the frequencies of transitions between word categories. Our criteria set comprises, all in all, 54 variables. Before employing them in Genesis, our team tested them in homogeneity problems of other Biblical books and found that they bring satisfactorily to light the finger prints of an author. And there, I wish to underscore, they invariably confirmed the views of the critical school against the traditional, much to the delight of the former.

The 20 000 words of Genesis were analyzed in great detail and each was marked with one of the three sigla J, E, P. On afterthoughts, we also added to each the sigla N, H or D, denoting whether the specific word appears in the mouth of the narrator (N) or in human (H) or in Divine (D) direct speech. With these data registered on magnetic tape, various statistical procedures were executed. Space permits to report on two only.

And now to results.

Par Wickmann extracted and pooled together the three corpuses J, E and P and calculated the probability of their origin in the same statistical population. He found the following :

PROBABILITY OF ORIGIN  
IN THE SAME STATISTICAL POPULATION

J	vs. E	82%
J	vs. P	0.000 0001 = 10 <sup>-7</sup> %
E	vs. P	0.000 0001 = 10 <sup>-7</sup> %

The cornerstone of the Documentary Hypothesis is here severely shaken : The comparison between J and E resulted in a homogeneity probability of 82 p.c. ! How enormous this percentage is may be seen in control tests applied to three German writers

CONTROL TESTS

	GOETHE	HERDER	KANT
GOETHE	22%	10-20%	10-42%
HERDER	-	71%	10-25%
KANT	-	-	9%

The probability of J and E being one is thus 10 times higher than that of Kant having written Kant and 4 times higher than that of Goethe having written Goethe.

To whether P must be considered a separate source I shall revert later.

Now, to the second approach, performed by H. Shore and myself. We operated with the matrix to be shown below.

	DOCUMENT J	DOCUMENT E	DOCUMENT P	TOTAL
words	5539	3294	2408	11241
NARRATOR N	NJ	NE	NP	N
samples	27	16	12	55
words	2849	3161	360	6370
HUMAN SPEECH H	HJ	HE	HP	H
samples	14	16	2	32
words	1021	161	611	1793
DIVINE SPEECH D	DJ	DE	DP	D
samples	5	1	3	9
words	9409	6616	3379	19404
TOTAL	J	E	P	TOTAL
samples	46	33	17	96

The columns represent the three Documents, the rows the three Sorts-of-Discourse. Thus, 9 cells obtained, unequally populated, of course. The 9 corpuses were divided into samples of about 200 words each. The first step was probing into the homogeneity of each cell which was soon established by Discriminant Analysis. After, Univariate, Multivariate, Discriminant, Factor, Reliability, Smallest Space and Cluster Analyses were performed. The final report, a voluminous book, is hopefully to be published. Here, I think it best to show a few visually easily perceivable results and round off this paper with a series of conclusions.

In the Discriminant Function Space, a computer-produced Scatter Diagram of all J, E, P samples looks as follows (fig. 1).

It is plain that, J- and E-samples form one compact configuration, their centroids almost coinciding, while P is remote. When the same approach was applied to the Discourse dimension, fig. 2 resulted.

Here, N-samples are a highly concentrated block, with those of H and D separate and interspersed with each other.

The distinction between the Jahwist and the Elohist is thus again not borne out. P, however, seems to be once more a source apart. Statistical Discourse Analysis shows an unmistakable caesura between N and H + D which cuts equally across all three Documents, as if these never existed.

Let us now proceed to Cluster Analysis. Discriminant Analysis takes as its point of departure the Documentary Hypothesis. Not so, though, Cluster Analysis. The program let the 96 samples form clusters or families, so to say by themselves, in accordance with the highest similarity between them on the basis of the criteria battery, fig. 3 is what emerged.

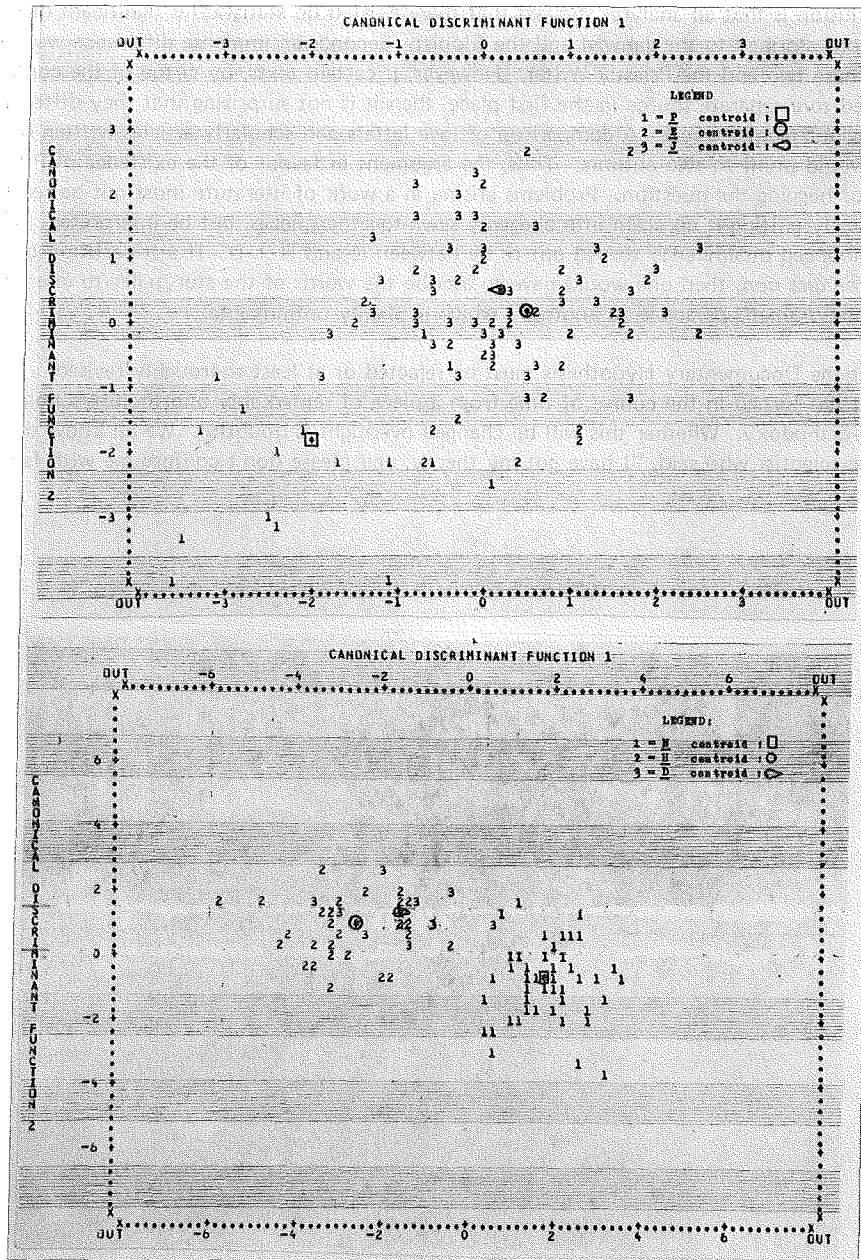
Samples are shown here as they group naturally disregarding any hypothesis, documentary or other. Their sequence is the same on the left-hand vertical column and on the horizontal bottom line so that the relationship between each pair of samples can be read by means of the coordinates. This relation, i.e. the proximity degree between two samples is indicated in shaded form : the darker the meeting point between the coordinates, the greater the proximity.

Let us now observe the diagram more closely. In the uppermost left-hand corner, there are three samples of highest proximity which have nothing to do with the rest. Since they are numbered and the cell into which they belong is given in the column 'Category', we can immediately identify them. They are of the NP cell, i.e. the Priestly Writer's as a narrator, and when looked up in the book found to be precisely the three long genealogies in Genesis. No wonder that they are isolated : they are of a different genre ! Then follows a large dark triangle. It consists without exception of N-samples with absolutely no distinction between J and E. In this way, the remainder should also be analyzed, but this would exceed the limits of this paper. At any rate, it is clear that the admittedly existing divergencies in language behaviour in Genesis are due to differences between Sorts-of-Discourse, a feature totally overlooked by Documentarians.

What are the conclusions to be drawn ? Before detailing them, I would like to qualify them. The Documentary Hypothesis refers to all the Five Books called Pentateuch whereas our enquiry deals solely with Genesis - though we are already working on Exodus. Further, that the conclusions, but not the data and results, may be questioned, although our conclusions were readily accepted by the critics when they suited their theories in previous investigations. Finally, that it was not our task to explain whence the inconsistencies and repetitions in Genesis. They can indeed be explained by conventional ways of literary criticism and are there for the asking, but not in the framework of this report.

The first conclusion is that all analyses showed that there existed no statistically significant differences between portions ascribed to the Jahwist and the Elohist. Second, an immense difference was evidenced between these two and the Priestly Writer. However, if certain texts, by virtue of the contents are apportioned to a separate source in the first place, then it is not surprising that they differ in language behaviour from the rest. A comparison of love letters and scholarly articles written by the same person would point to two authors. Thus, the argument in favour of the existence of P is a classical case of begging the question. Problems arising in a work of literature must not be resolved by mechanistically cutting it up again into allegedly sewn-together pieces, but be interpreted. Third, the most pronounced dissimilarity turned out to be between N and H + D. If a tripartition of Genesis is a *sine qua non*, then one would have to ascribe the words of the storyteller to one hand and those of the *dramatis personae* to another - which is plainly unthinkable.

It follows that the Documentary Hypothesis must be rejected or at least thoroughly revised. Regrettably, it has turned in the course of time from a piece of remarkably unorthodox scholarship into scholarly orthodoxy. Whether this will be changed overnight is doubtful. We all know the attitude of the scientist who said, "I have got my theory, and please don't confuse me with facts."



No. CPT. DM  
 1 EP I I  
 2 EP I I  
 3 EP I I  
 4 EP I I  
 19 NJ I I  
 33 NJ I I  
 5 EP I I  
 20 NE I I  
 27 NE I I  
 24 NE I I  
 15 NE I I  
 18 NE I I  
 28 NE I I  
 25 NE I I  
 17 NE I I  
 30 NJ I I  
 44 NJ I I  
 19 NE I I  
 53 NJ I I  
 46 NJ I I  
 47 NJ I I  
 52 NJ I I  
 25 NE I I  
 14 NE I I  
 40 NJ I I  
 51 DJ I I  
 53 DJ I I  
 71 NE I I  
 92 DJ I I  
 86 DJ I I  
 91 DJ I I  
 74 DJ I I  
 61 DJ I I  
 81 DJ I I  
 79 DJ I I  
 84 DJ I I  
 59 NE I I  
 48 NJ I I  
 72 NE I I  
 47 DJ I I  
 35 DJ I I  
 57 DJ I I  
 50 DJ I I  
 6 EP I I  
 32 DJ I I  
 99 DJ I I  
 66 DJ I I  
 89 DJ I I  
 36 DJ I I  
 23 DJ I I  
 42 DJ I I  
 69 DJ I I  
 93 DJ I I  
 34 DJ I I  
 64 DJ I I  
 75 DJ I I  
 74 DJ I I  
 26 NJ I I  
 31 DJ I I  
 30 DJ I I  
 25 DJ I I  
 22 DJ I I  
 54 DJ I I  
 35 DJ I I  
 84 DJ I I  
 62 DJ I I  
 29 DJ I I  
 86 DJ I I  
 21 DJ I I  
 85 DJ I I  
 80 DJ I I  
 68 DJ I I  
 94 DJ I I  
 7 NP I I  
 9 NP I I  
 12 NP I I  
 11 NP I I  
 3 NP I I  
 91 DJ I I  
 67 NP I I  
 20 NP I I

THE DISTANCES ARE REPRESENTED IN SHADED FORM ACCORDING TO THE FOLLOWING SCHEME:

	FROM	LESS THAN	TO	PROXIMITY DEGREE
M	7.005	7.005	7.984	8
M	7.754	7.984	8.337	7
X	8.337	8.337	8.856	6
.	8.856	8.856	9.405	5
.	9.405	9.405	10.232	4
.	10.232	10.232	11.232	3
.	11.232	11.232	GREATER THAN	2

EXAMPLE: THE DISTANCE BETWEEN SAMPLE NO. 69 AND SAMPLE NO. 96 IS LESS THAN 7.005

Sample no. 33 5222411243134154445211 445967987 977 66877854758355 3968912 4699677723 362595062828869 13 864  
 12493550719588577849367325431 016 31267185 10390298 247570 862069663 2 934534610 5285642981508479213170