

The Nations Sample

By Doron Witztum

Abstract:

In 1995 I published (as a preprint, together with Professor Eliyahu Rips and Yoav Rosenberg) an article entitled: "Equidistant Letter Sequences in the Book of Genesis: II. The Relationship to the Text" [1]. This article dealt with convergences between expressions appearing as ELS's (Equidistant Letter Sequences) and expressions appearing in consecutive letters in Genesis. One of the samples discussed in the article was the Nations Sample (for a description of this sample in Hebrew see [2]). Measurements conducted on this sample indicated a particularly high level of statistical significance. In light of criticisms leveled against the composition of this sample and its measurement, in an article [3] by Dr. D. Bar Natan, Dr. B. McKay and Prof. S. Sternberg (henceforth, BMS), I am presenting here a description of the considerations and decisions that went into the construction of the sample. I will also present a new measurement of the significance. The criticisms of BMS are refuted in this article, and careful analysis of their data and suggestions leads to new results supporting our research hypothesis with high significance.

The structure of the article:

BMS distinguished between two parts of the Nations Sample. The simpler part of the sample they label "regular," and the other part they call "irregular". In the present article I will focus on the part of the sample that they considered simpler - the "regular" part. I will deal with it in two parts:

Part I: "Possible Presentations of the chosen Topic" – This section will deal with the various ways of compiling the regular part of the sample for the chosen topic.

Part II: "The Range of Stories" – In this section I will discuss the range of possibilities in choosing a topic for such a sample.

(A discussion of the "irregular" part of the sample is the subject of a planned Part III.)

Part I: Possible Presentations of the Chosen Topic

Introduction:

In this introduction I will describe briefly how the Nations Sample was compiled. For the sake of clarity I will not address the criticisms of BMS at this point (I will do so in the chapters to come).

A. In chapter 10 of Genesis there is a list of 70 names of descendants of Noah: the children of Shem, the children of Ham and the children of Japhet. Jewish tradition [4] treats these as the forebears of the 70 nations from which all mankind are descended. Later scholars also take this as a given, and this list of nations can be found in various Biblical encyclopedias under the heading "The Table of Nations" or "The Children of Noah" [5].

I wanted to investigate whether the fact that Noah's 70 descendants developed into the 70 nations, is encoded in the Book of Genesis. Specifically, I wanted to see

whether this can be demonstrated through convergences between the names of Noah's descendants and encoded expressions reflecting their eventual nationhood. It must be clear that the expressions indicate the names of nations. This can be achieved by first identifying the kinds of terms that indicate nationhood.

B. One classical approach maintains that there are objective criteria by which nations can be distinguished from one another. This fundamental issue is discussed in the *Encyclopedia Hebraica* [6]:

“Much thought has been given to the question: What are the indicators of nationhood, and how it is to be determined whether a particular individual belongs to one nation or another? There are two fundamental approaches to this subject. One approach looks for the answer to this question in objective criteria. According to this approach a nation is said to exist as an independent entity if it consists of a group of people who share a common language (שפה), are distinctive in their cultural forms and way of life, are concentrated in a certain territory, and frequently share a common religion as well. One extreme opinion holds that they must also be of common racial stock. According to this approach an individual's membership in a national group is determined by investigating these criteria.”

We see that the *Encyclopedia Hebraica* recognizes three fundamental signs of nationhood:

1. Language (שפה) [7]. 2. Culture and way of life. 3. Territorial concentration (land). These are frequently accompanied by a fourth indicator, religion, and according to an “extreme opinion,” a fifth indicator, racial stock.

In the *Concise Oxford Dictionary* [8], under the entry “nation,” I found the following definition: “A distinct race or people having common descent, language, history, or political institutions.”

Other dictionaries and encyclopedias [9] give perhaps a few other indicators, but in any event there are altogether only a handful of them, and each of the sources emphasizes the ones it thinks are most essential.

C. In traditional Jewish literature I found a discussion of this subject in the following excerpt from the Vilna Gaon [10] (for a comparison of this excerpt with quotes from the Vilna Gaon's other writings see chapter V):

“It is well known that the nations mentioned in the Holy Scriptures were distinguished from one another by four characteristics:

1. By their names. Each nation (עם) was called by the name of their first ancestor from whom they were descended, such as Cush (כוש), Mitzraim (מצרים), Put (פוט), Canaan (כנען), and the like from among the nations, and their descendants are all called Mitzrites and Edomites, after their first father, to this very day.
2. By the names of their countries (מדינות). Countries are also named after their founders. The Land of Mitzraim (ארץ מצרים), for example, is called after its founder. So, too, the Land of Cush (ארץ כוש) and the Land of Canaan (ארץ כנען), etc.
3. By language (לשון). The members of each nation have agreed upon a distinctive language, called the Mitzrite language (לשון מצרי), the Cushite language (לשון כוש), etc.
4. By script (כתב). Each nation has a distinctive script.”

From this passage we can learn about four indicators of nationhood:

1. Name. 2. Land. 3. Language. 4. Script.

D. The regular part of the Nations Sample:

We must now look for four expressions that reflect the four indicators mentioned in **C**. In the regular part of the sample we will search for expressions of the form “**prefix** X,” where X is one of the names from the “Table of Nations” and the “**prefix**” is a word in its appropriate Hebrew form, i.e., the construct case [*semichut*]. It is quite possible that each of these indicators can be expressed in more than one way. Therefore it must be decided ahead of time how the expressions will be selected.

In view of our research hypothesis (that there is a hidden text encoded in the book of Genesis) it seems reasonable to assume that when Biblical themes are encoded preference will be given to the language used by the Scriptures themselves. For this reason my approach was to look for those combinations that are closer to the language of the Scriptures. How this was carried out in practice is described in Chapters III and IV. Here I will just list the expressions that were settled upon:

1. "עם X". 2. "ארץ X". 3. "שפת X". 4. "כתב X".

For example: If X is the name גמר, the first name on the “Table of Nations” (found in the Appendix), we will receive for it the expressions: עם גמר, ארץ גמר, שפת גמר, and כתב גמר. We will then measure convergences between the name גמר, as it appears in consecutive letters (that is, with a skip length of $d = \pm 1$) and each of the four expressions as they appears as ELS's. In other words, we will investigate convergences for each of the word pairs: (גמר, עם גמר), (גמר, ארץ גמר), (גמר, שפת גמר) and (גמר, כתב גמר). We will do this for every name X appearing on the “Table of Nations” [11]. The set of word pairs obtained in this way constitutes the regular part of the Nations Sample.

E. Results of the regular part of the sample:

1. When we measure the significance of the sample using the original randomization test [1] we receive the following rankings:

- The rank of P_1 out of 1,000,000,000 values of P'_1 was 70. That is,
 $r_1 = 7.0 \times 10^{-8}$.
- The rank of P_2 out of 1,000,000,000 values of P'_2 was 5667. That is,
 $r_2 = 5.67 \times 10^{-6}$.

Thus the overall significance using this method was $r = 2 \times r_1 = \underline{1.4 \times 10^{-7}}$.

2. In our article [2] we discussed samples similar the Nations Sample in addition to samples of a different type, for the sake of which a different randomization test (henceforth, *RPWL* = Randomization by Permutations of Words' Letters) was used. At a later date the *RPWL* test was applied to samples like the Nations Sample as well [12]. In their article, BMS criticized the use of the original test for the Nations Sample, claiming that the results obtained thereby were meaningless for a number of reasons particular to this instance [13].

Since these criticisms are not relevant to the *RPWL* test, it is preferable to use this test for measuring the significance. Here are the results:

- The rank of P_1 out of 1,000,000,000 values of P'_1 was 14.
That is, $r_1 = 1.4 \times 10^{-8}$.

- The rank of P_2 out of 1,000,000,000 values of P'_2 was 2,773.
That is, $r_2 = 2.77 \times 10^{-7}$.

Thus the overall significance using this method was $r = 2 \times r_1 = \underline{2.8 \times 10^{-8}}$.
For more details of the measurements see the Appendix.

I. Why עַם?

A1. BMS [14] proposed that the statement of the Vilna Gaon that “Every nation is called after their first father” should also be included in the experiment by looking for expressions of the form “X שם” (“name of X”). In other words, we should investigate whether or not the expression שם כנען (name of Canaan) converges with the name כנען (Canaan) in the text.

Response:

This is simply a misreading of the words of the Vilna Gaon. The Gaon was addressing the question: “By what name is a nation called?” BMS are in effect proposing that we investigate the question itself! It is almost like the famous joke where one person says to another, “Please tell me - what is your name?” to which his companion responds by telling him literally the words, “What is your name.”

A2. BMS further proposed investigating whether the name כנען in ELS form converges with the name כנען appearing in consecutive letters in the text.

Response:

I will try to illustrate the flaw in this proposal with the help of the following example: When we open up the dictionary [9] to the entry “ישראל” (Israel) we discover that this is a masculine personal name. In addition, however, we find the combinations “עם ישראל” (the nation of Israel), “ארץ ישראל” (the land of Israel), and “מדינת ישראל” (the state of Israel). We also learn that the name “ישראל” alone can be used in place of all three of these combinations.

When the name “ישראל” is used it can usually be deduced from context whether what is meant is the Jewish people, the land of Israel, the state of Israel, the patriarch Jacob or a child whose name happens to be “ישראל.” On the other hand, when one considers the name “ישראל” in isolation, not in a context, how can one determine which of these possibilities is intended?

What kind of message can be encoded by BMS’s proposed tautologies? It gives us no new information, and it certainly does not indicate to us whether each of the descendants of Noah became a nation in his own right! The research hypothesis itself predicts that both of BMS’s proposals should fail, and indeed (as BMS report) both of them do fail. Their assertion that these two proposals form “the most reasonable constructions suggested by the Vilna Gaon’s explanation” is simply astonishing

B. BMS further asserted [14] that we combined the prefix עַם with a plural form of the name, as in the case of לודים (Ludim), and that such a combination is totally unjustified by the Vilna Gaon’s first criterion.

Response:

1. First of all, the reader should be informed [15] that despite the apparent "ים" suffix, **לודים** is actually the name of an individual, as are the names **ענמים**, etc. Therefore there is no difference between this name and the names **גמר** or **יון**.
2. Even according to those commentaries who do say that **לודים** is a plural name, and not the name of an individual, nevertheless they agree that this is the name by which the nation was known. Therefore it is entirely legitimate to use the combination **עם לודים** [16].

C. BMS [14] replaced the word **עם** with the word **גוי**, claiming that “the Torah uses the word **גוי** for nation, not **עם**.” They justify this claim by citing a passage from a commentary by the Vilna Gaon (*Adereth Eliyahu* on Isaiah 1:4).

Response:

Both their assertion and evidence are incorrect.

1. The assertion that the Torah uses only the word **גוי** for nation is simply false. In the *Encyclopedia Biblica* (at the entry **עם**, v. 6, p. 235) we find a discussion of Scriptures usage of the term **עם**:

“The use of terms for a nation or people, in all their various shades of meaning, is the most frequent in the language of the Scriptures. The same connotations are indicated also by the words **גוי** and **לאום**. There are some (L. Rost, E. A. Speiser) who wish to distinguish between the Biblical uses of the terms **עם** and **גוי** in the following manner: **גוי** indicates the citizens of a certain kingdom or geographic region, whereas **עם** (according to Speiser) indicates common origin, and has overtones of a familial relationship, a shared destiny and a relationship with the national god.”

In other words, not only is the term **עם**, meaning nation or people, widespread throughout the Bible, but in this instance it is a more appropriate term than **גוי** because it emphasizes common origin.

The reader should note, however, that this was not the basis for my decision to use **עם** rather than **גוי** (for a discussion of this decision see Ch. IV). In any event, it is quite clear that the assertion that the Torah uses only the word **גוי** for nation is simply incorrect.

2. The evidence BMS cite does nothing to enhance their assertion. The relevant passage from the Vilna Gaon (*Adereth Eliyahu* on Isaiah 1:4) reads as follows:

“The difference between **גוי** and **עם** is that **עם** is used to refer to a collection of many individuals, whether of the masses, whether possessed of a religion or not. **גוי**, on the other hand, only refers to those who properly adhere to custom. Even an individual who observes the customs received by the multitude is called a **גוי**. This is what the Sages mean when they say that [the Israelites in Egypt] ‘became there a **גוי** – indicating that Israel behaved distinctively there’.”

In other words, the term **עם** is correct in any context. Only the term **גוי** is limited to specific contexts. Furthermore, **גוי** can sometimes refer to a single individual.

It is worth noting that in two other places (Habakkuk 2:5 and I Chronicles 16:24) the Vilna Gaon writes exactly the opposite – that **גוי** indicates the masses. In any event, as mentioned above, I did not base my decisions upon such distinctions. I

took from the Vilna Gaon only the four indicators of nationhood, not the terms by which they are designated.

For anyone who still believes that the term גי was no less appropriate a choice than the term ע there exists the simple option of calculating the significance of the sample twice:

a) Once using ע as the first term, and

b) once using גי as the first term.

To account for this supposed freedom all one need do is multiply the best significance by a factor of 2

II. לשון Versus שפה

In their article [14] BMS note that the Vilna Gaon, in the passage cited in the Introduction, used the term לשון for “language,” whereas when I compiled the sample I used the term שפה. BMS translate the term לשון as “dialect” and the term שפה as “language.” Their justification for this interpretation is the commentary of Rabbi Samson Raphael Hirsch [17] (on Genesis ch. 10). On the basis of this interpretation BMS claim that when the Vilna Gaon uses the term לשון he means “dialect,” whereas I used “language” as the indicator of nationhood.

Response:

- A. A dialect is not an indication of nationhood; language is (as I demonstrated in the introduction with the help of the *Encyclopedia Hebraica* and the *Oxford Dictionary*).
- B. It is interesting to note that according to the very source cited by BMS, the commentary of Rabbi S. R. Hirsch, the indicator of nationhood is שפה (“language”) not לשון (“dialect”) [18] – precisely confirming my own word choice! It is strange that BMS neglected to point this out.
- C. Now let us deal with the words of the Vilna Gaon himself. The Hebrew language has many layers [19], beginning with the language of the Scriptures right up to the Hebrew used today. The Vilna Gaon was writing in the language of the rabbis (לשון חכמים). The Sages have a general principle: “The language of the Torah is one thing and the language of the rabbis is another” [20]. The term לשון is used by the rabbis as a synonym for the word שפה appearing in the Bible. This is particularly true regarding its usage in the context of the descendants of Noah, as can be seen from [21]. Thus the assertion of BMS that the Vilna Gaon meant specifically “dialect” when he used the word לשון is without foundation.

Summary: It turns out that even according to the commentary cited by BMS the correct choice was שפה. In the chapters to come it will be explained how this word choice was actually arrived at through a “mechanical” procedure (i.e., one involving no interpretation).

For anyone who still believes that the term לשון was no less appropriate a choice than the term שפה there exists the simple option of calculating the significance of the sample twice:

a) Once using שפה as the first term, and

b) once using לשון as the first term.

To account for this supposed freedom all one need do is multiply the best significance by a factor of 2.

N.B.: According to Rabbi Hirsch, a single language may have several לשונות (dialects). Therefore if BMS considered use of the concept “dialect” to be so crucial, they should have used לשונות (dialects) rather than לשון.

III. The Range of Possibilities:

The “regular” part of the sample consists of expressions of the form “**prefix** X,” where X is one of the names from the “Table of Nations,” and the “**prefix**” is a word in the construct case [*semichut*]. Given our research hypothesis, it seems reasonable to expect that when Biblical themes are encoded there will be a preference for expressing them in the language of the Scriptures themselves.

In this chapter and the next we will explore ways of compiling a list of Scriptural prefixes to indicate the four categories of the Vilna Gaon:

1. The name of the nation. 2. Its land. 3. Its language. 4. Its script.

We will begin by looking at the words used by the Gaon himself: מדינה, עם, אמה, גוי, כתב, לשון, ארץ.

A. First we must examine whether these terms are indeed found in the Bible or not. We will locate each through use of the *New Concordance of the Bible* [22]. This concordance lists occurrences of each word according to meaning. The meaning is defined by a synonym or defining phrase.

In this way I was able to identify the Scriptural synonyms for each term (for full details see the Appendix). Here is a list of the possibilities arranged by category:

1. גוי, לאם, אמה, עם.
2. מדינה, ארץ.
3. שפה, לשון.
4. כתב.

B. The fact that a word occurs in the Bible does not guarantee that Scripture uses it in the construct case. We must again use the concordance to determine which of the terms in **A.** occur in the desired form (the *New Concordance* lists this form separately). It turns out that the following words occur in this form (i.e., singular construct case):

1. עם, גוי. 2. ארץ. 3. שפה, לשון. 4. כתב.

Summary:

There are **two** possibilities for representing category 1, **one** for category 2, **two** for category 3, and **one** for category 4. In all there are $1 \times 2 \times 1 \times 2 = 4$ possibilities. There is another possibility as well – to use all of them. Thus there are a total of **5** ways of compiling the regular part of the Nations Sample.

Although in our original experiment on this sample we subjected this list to a further selection process, as described in the next section, for the sake of the reader’s curiosity I will present here the results received when of all the terms are used (that is, עם as well as גוי, and שפה as well as לשון):

Using the original randomization test, with 1,000,000,000 permutations, we receive:

$$r_1 = 1.48 \times 10^{-6}, \quad r_2 = 3.70 \times 10^{-6}$$

Using the *RPWL* test, with 1,000,000,000 permutations, we receive:

$$r_1 = 3.21 \times 10^{-7}, \quad r_2 = 1.37 \times 10^{-6}$$

IV. Decisions:

In the previous chapter I listed five ways of compiling a list of Scriptural prefixes for the regular part of the Nations Sample. This is the number of ways we can conduct our experiment if we choose to ignore all other considerations, that is, if we do not try to distinguish between עַם and גּוֹי (in category 1) and between לִשׁוֹן and שִׁפְהָ (in category 3).

However, we must explore the possibility that there are differences in the usage of these synonyms in Scripture. Some of them might turn out to be more “appropriate” than others for use in a combination of the form “prefix X” (where X is the name of a nation). This question can be addressed with through interpretation, as we discussed in Chapter I regarding the terms עַם and גּוֹי, and in Chapter II regarding לִשׁוֹן and שִׁפְהָ. Here, however, we will try to deal with it in a “mechanical” fashion to avoid the need for interpretation.

Decision by frequency:

About a year after we first published our work on the Nations Sample [1], Prof. Havlin’s Report was published [23], in which he described the guidelines by which he compiled the list of famous rabbis. It turns out that in certain cases he distinguished between various alternatives based on their frequency in the database [24].

Taking advantage of the fact that the New Concordance provides a list of all instances of construct case, we will check the frequency of the desired form (singular construct case) for each of the words in categories 1 and 3 (where there was more than one alternative). Then we can make a selection based on frequency.

1.

Word	Singular, Construct Form	Number of Occurrences
עַם	עַם	63
לֵאם	לֵאם	0
אִמָּה	אִמָּה	0
גּוֹי	גּוֹי [25]	1

3.

Word	Singular, Construct Form	Number of Occurrences
לִשׁוֹן	לִשׁוֹן	3
שִׁפְהָ	שִׁפְהָ	5

Results: Using X to indicate one of the names on the Table of Nations, the expressions we will look for are:

1. X עַם. 2. X אֶרֶץ. 3. X שִׁפְהָ. 4. X כְּתָב.

N.B.: In the case of עַם and גּוֹי the decision was clear. In the case of שִׁפְהָ, however, there is room to doubt whether we should base our selection upon a simple majority of usage, or should we only select the more frequently used term when there is an order of magnitude between them.

Decision by “precedent”:

This is the procedure I actually used in the original experiment.

For each of categories 1 and 3 (the ones for which there are multiple possibilities) I selected the prefix for which there was a precedent in the Scriptures:

Using the Concordance, I investigated whether a particular prefix appears in the Biblical text itself in the desired combination, “prefix X,” with one or more of the names from the “Table of Nations.”

The fact that a term is used in the text itself with the name of a nation makes it more certain that it is indeed an indicator of nationhood. The reason for insisting that it appear specifically in combination with one of the 70 nations is that only for these nations is there a tradition [4] that they constitute separate nations in every respect. If a term appears in combination with a name that is not on the Table of Nations there is no guarantee that it is in fact an indicator of nationhood, since there is no way of knowing whether the name it appears with is that of a nation [26].

Results: As it turns out, for the first category only **עם**, and for the third category only **שפת** appears in the Biblical text in the desired combination. In other words, again we receive:

1. X **עם**. 2. X **ארץ**. 3. X **שפת**. 4. X **כתב**.

V. Why Did We Follow this Particular Passage from the Vilna Gaon?

In their critique [27] BMS note that they found two other places where the Vilna Gaon deals with our subject. For the reader’s benefit I will cite these two other sources:

1. One is in the Gaon’s commentary *Adereth Eliyahu* on Isaiah 1:4:
 “After the flood the people divided into seventy nations, whether because of their division into separate families (משפחות), or because of separated languages (לשוניות), or separated lands (ארצות), or separate manners (גימוסים). This is what is meant by “to their families, in their lands, in their languages, in their nationalities.”
2. The other source is the Gaon’s non-Kabbalistic commentary to Esther 1:22:
 “Because nations are divided by four things: script, language, land – in that each nation has its own land, and garments.”

The reader is invited to compare these two excerpts with the one from Job that I cited in the Introduction. In my opinion it is clear that only the excerpt from Job contains a thorough treatment of the indicators of nationhood. Regarding the excerpt from Isaiah in particular, it seems clear that the Gaon is merely discussing the causes of the division into nations, not the indicators of nationhood.

If the reader is not interested in making such distinctions in the wording of the Gaon, we recommend that he employ the simple strategy of multiplying the significance by the number of options: There are three ways to choose one of the three texts of the Gaon, and there is another option to take the union of all of them. Thus there are four options altogether.

Revising the number of possibilities:

If one chooses to ignore all interpretation and avoid making any decision, the choice of the Gaon’s commentary to Job is one of four possibilities, and the selection of the four prefixes used in the experiment to describe the categories of the Gaon is one of five possibilities. Thus there are $4 \times 5 = 20$ possibilities in all.

For the sake of the reader who is curious, I will present here the results for two of these options:

A. The option of using all three excerpts from the Gaon adds another three prefixes found in Scripture in the appropriate construct form: **משפחת**, **בגדי**, **מנהג** (see the Appendix for details).

- Using the original randomization test, with 1,000,000,000 permutations, we receive:

$$r_1 = 7.0 \times 10^{-9}, \quad r_2 = 7.58 \times 10^{-7}$$

- Using the *RPWL* test, with 1,000,000,000 permutations, we receive:

$$r_1 = 2.0 \times 10^{-9}, \quad r_2 = 4.23 \times 10^{-7}$$

B. If we calculate the significance of the combination of the three excerpts, without distinguishing between alternative prefixes, provided that all are found in Scripture in the appropriate construct form (this is the option mentioned in Chapter III in which we use both **עם** as well as **גוי**, and **שפה** as well as **לשון**), we receive:

- Using the original randomization test, with 1,000,000,000 permutations, we receive:

$$r_1 = 1.15 \times 10^{-7}, \quad r_2 = 5.26 \times 10^{-7}$$

- Using the *RPWL* test, with 1,000,000,000 permutations, we receive:

$$r_1 = 2.6 \times 10^{-8}, \quad r_2 = 2.26 \times 10^{-7}$$

Appendix to Part I

The Table of Nations:

Here is a list of the names of the 70 descendants of Noah as they appear in Genesis, chapter 10 [compare with the entry **לוח העמים** (The Table of Nations) in *Encyclopedia Biblica*, v. 4, pp. 443-444]:

1. גמר, 2. מגוג, 3. מדי, 4. יון, 5. תבל, 6. מוש, 7. תירס, 8. אשכנז, 9. ריפת, 10. תגרמה, 11. אלישה, 12. תרשיש, 13. כתי, 14. דדני, 15. כוש, 16. מצרים, 17. פוט, 18. כנען, 19. סבא, 20. חוילה, 21. סבתה, 22. רעמה, 23. סבתא, 24. שבא, 25. דדן, 26. לודים, 27. ענמים, 28. להבים, 29. נפתחים, 30. פתרים, 31. כסלחים, 32. פלשתים, 33. כפתרים, 34. צידן, 35. חת, 36. היבוסים, 37. האמרי, 38. הגרגשי, 39. החוי, 40. הערקי, 41. הסיני, 42. הארודי, 43. הצמרי, 44. החמתי, 45. עילם, 46. אשור, 47. ארפכשד, 48. לוד, 49. ארם, 50. עוז, 51. חול, 52. גתר, 53. מוש, 54. שלח, 55. עבר, 56. פלג, 57. יקטן, 58. אלמודד, 59. שלף, 60. חצרמות, 61. ירח, 62. הדורם, 63. אוזל, 64. דקלה, 65. עובל, 66. אבימאל, 67. שבא, 68. אופר, 69. חוילה, 70. יובב.

Note that two of the names are listed twice: **שבא** and **חוילה**. For the sake of our measurements they were counted only once each. In other words, we actually only used 68 different names.

The Range of Possibilities:

Locating synonyms in the Scriptures:

Let us begin with the terms that the Vilna Gaon used: **עם**, **מדינה**, **ארץ**, **לשון**, **כתב**, **גוי**, **אמה**. Let us investigate whether these words are used in Scripture, and let us identify Scriptural synonyms for each one with the help of the *New Concordance of the Bible*.

In the *Concordance* occurrences of each word are organized according to meaning, and the meanings are defined by a synonym and/or defining phrase.

Locating Scriptural synonyms:

- 1a.
 - a) Under the entry **עם** (p. 882) we find 5 different shades of meaning. The first is the desired one. It has the following synonyms: **אמה, לאום, גוי**.
 - b) Next let us look up each of these synonyms in the *Concordance* to see if they, too, are used in Scripture, and to verify their spelling according to the entry heading. After we do this we are still left with: **אמה, לאם, גוי**.
- 1b.
 - a) At the entry **גוי** (p. 228) we find two meanings listed. The first is the desired one. We find the following synonyms: **עם, אמה**.
 - b) We have already discovered above that both of these are found in Scripture and this is how they are spelled there.
- 1c.
 - a) At the entry **אמה** (p. 82) we find only one meaning, for which there are two synonyms: **עם**, and **שבט**.
 - b) The word **שבט** has a different connotation [“tribe”], making it unsuitable for use as an indicator of nationhood. [This can be verified by checking the entry for **שבט** in the *Concordance* (p. 1104) and noting that it does not have listed as synonyms words such as **עם, אמה, לאם** or **גוי**, or anything similar to them.]
 - c) Thus we are left with the word **עם** as the only synonym for **אמה**.
- 2a.
 - a) At the entry **ארץ** (p. 112) we find three different shades of meaning. The third one is the desired one. One synonym is listed: **מדינה**.
 - b) Using the *Concordance* we investigate whether this synonym is used in Scripture and how it is to be spelled. The result is that we are left with the word: **מדינה** alone.
- 2b.
 - a) At the entry **מדינה** (p. 623) we find only one meaning, for which there is one synonym: **ארץ**.
 - b) We have already discovered above that **ארץ** is Scriptural, and this is its correct spelling.
3.
 - a) At the entry **לשון** (p. 611) we find four different meanings. The second one is the desired one. The synonyms listed are: **שפה** and **דבור**.
 - b) Actually, the connotation of the word **דבור** [speech] makes it unsuitable as an indicator of nationhood, although the *Concordance* does not make this distinction, but in any event this word does not appear in Scripture, therefore we are left with: **שפה**.
4.
 - a) At the entry **כתב** (p. 567) we find two meanings. The second one is the desired one. No synonyms are listed.

Prefixes found in the other excerpts from the Vilna Gaon:

A. The prefixes found in *Adereth Eliyahu* on Isaiah 1:4:

We have already seen the actual citation, see Chs. V and I.

BMS proposed that we examine the prefixes:

1. משפחת. 2. לשון. 3. ארץ. 4. נמוס.

1) משפחת – Actually, it is not implied in the wording of the Gaon that this is an indicator of nationhood. Nevertheless, we will go ahead and identify its Scriptural synonyms:

a) In the *Concordance* at the entry משפחה (p. 721) we find three definitions, the second of which is the desired one. We find listed there the one Scriptural synonym: לאום.

b) Again using the *Concordance* we check the spelling of the synonym לאום as it appears in the Bible, and we find that it should be written לאם.

c) Of these two choices only the word משפחה appears in the Scriptures in construct case: משפחת.

2) לשון – This has already been dealt with above.

3) ארץ – This has already been dealt with above.

4) נמוס – This word does not appear in the *Concordance* because it is not a Biblical word. The algorithm described in Ch. III (“The Range of Possibilities”) does not allow for such cases. One possibility is to simply erase it. BMS preferred to use it, therefore I will do the same. I opened the *New Dictionary* [9] (by the same author as the *New Concordance*) and discovered that the proper spelling of this word is נימוס (which is how it is spelled in the excerpt from the Gaon as well). The dictionary lists one synonym: מנהג. This word does appear in Scripture, and in the desired form – singular construct [28]. While it is true that some Biblical commentaries render this word as “driving,” nevertheless, there are those [29] who interpret it as “manners” (contrary to the words of BMS). Therefore מנהג is certainly preferable to נימוס (which all agree is not a Biblical usage).

Summary: According to this source we have the following possible prefixes:

1. משפחת. 2. לשון, שפת. 3. ארץ. 4. מנהג.

If we use the same selection process described earlier in the article to distinguish between the two choices in category 2, we receive:

1. משפחת. 2. שפת. 3. ארץ. 4. מנהג.

B. The prefixes found in the commentary to Esther 1:22:

For the actual citation see Ch. V. BMS proposed using the prefixes:

1. כתב. 2. לשון. 3. ארץ. 4. מלבוש, לבוש, בגד.

The words for categories 1-3 have already been dealt with above. All that remains for us to deal with is category 4:

מלבוש – Here BMS made a peculiar error; for some reason they wrote this word in the singular, although the Gaon uses it in the plural, מלבושים. In the English, too, they translated it “dress” rather than in the plural, “garments.”

1. When we examine the entry for the word מלבוש (p. 661) in the *Concordance* we find only one meaning. Two synonyms are listed: לבוש and בגד.

2. With the help of the *Concordance* we learn that both of these words are Biblical, and that this is their proper spelling.

3. Of the three choices, only בגד appears in the Scriptures in the appropriate construct form: בגדי.

Summary: From this source we receive the following possibilities:

1. כתב. 2. לשון, שפת. 3. ארץ. 4. בגדי.

If we apply to category 2 the same selection process as before, we receive:

1. כתב. 2. שפת. 3. ארץ. 4. בגדי.

The measurement:

A. The Analysis method used in [1]:

The Analysis method used in [1] is essentially the same as in the Statistical Science paper [30], except that the names of the nations are sought at skip +1 and -1 only. The related phrases are still sought as ELSs. The definition of $c(w, w')$ for the present case is given in the Appendix of [1], Sections A.1 and A.2. To read it [click here](#).

B. The RPWL Test:

The RPWL test is explained in [2] and [12]. Here we applied it as follows:

1. We took one of the Nations Sample's pairs, (X, prefix X), and carried out 1000 permutations (including the identity permutation) of the expression "prefix X". In the event that the number of possible different permutations n was less than 1000, we performed n permutations. The permutations were conducted in a standardized manner using a program designed by Yaakov Rosenberg.

For example: the first pair is (גמר, עם גמר). We shall present here some of the pairs which are formed by the permutations (by order, from left to right):

גמר עמגמר	גמר גערממ	גמר געמרמ	גמר געממר	...	גמר ממעגר	גמר ממרגע
--------------	--------------	--------------	--------------	-----	--------------	--------------

2. We calculated the values of $c(w, w')$ for the convergences of all 1000 (or n) permutations of "prefix X" as ELSs, with "X" taken as it appears in consecutive letters (that is, with a skip length of $d = \pm 1$). For example, with regards to the example above, we obtain a row of cells. In each cell there is a c -value of the specific pair:

92/125	34/125	95/125	98/125	...	21/125	34/125
--------	--------	--------	--------	-----	--------	--------

In a case that a certain permutation of "prefix X" did not appear as ELS, it resulted with an empty cell.

3. Stages 1 and 2 were performed with regards to all the pairs in the Sample. We thus obtained rows of cells, each containing 1000 (or n) cells. In each cell which is not empty, there is a c -value of the convergence (X, prefix X).

4. We then chose by lottery one of the cells in the first row, one of the cells in the second row, and so on. We obtained a set of values of $c(w, w')$ and we calculated the values of P_i for them.

5. We repeated this procedure 999,999,999 times, using an algorithm for randomization similar to that described in [1]. The program used was also prepared by Yaakov Rosenberg. We used the same seed as in [1] and [30].

Bibliography and Notes:

1. D. Witztum, E. Rips and Y. Rosenberg, *Equidistant Letter Sequences in the Book of Genesis: II. The Relation to the Text*, a preprint (1995).
2. D. Witztum, E. Rips and Y. Rosenberg, *A Hidden Code in Equidistant Letters in the Book of Genesis: The Statistical Significance of the Phenomenon*, a preprint, spring 1996.
3. D. Bar Natan, B. McKay and S. Sternberg, "On the Witztum-Rips-Rosenberg Sample of Nations", preprint, March 11, 1998.
4. A source for this tradition can be found in *Targum Yonathan* on Deuteronomy 32:8.
5. See, for example, *Encyclopedia Biblica*, Bialik Institute, Jerusalem 1962, at the entry לוח העמים ["Table of Nations"] (v. 4, p. 439).
6. *Encyclopedia Hebraica*, Encyclopedia Publishing Company, Ltd., Jerusalem, 1981, at the entry לאום, לאומיות ["nation, nationhood"] (v. 21, p. 66).
7. This is clarified in the continuation of the words of the encyclopedia: "The difficulty with this approach is threefold: First of all, it is difficult to establish objectively the distinctiveness of dialects, etc." In other words, it is difficult to determine whether the language spoken by a group of people constitutes a separate language, indicating their distinctive nationhood, or whether it is merely a dialect of a more generally spoken language.
8. *The Concise Oxford Dictionary of Current English*, Oxford, at the Clarendon Press, 1972.
9. For example, A. Even Shoshan, *A New Dictionary of the Hebrew Language*, Kiryat Sefer, Jerusalem, 1989, at the entry עם ["Nation"], (p. 988):
 "A large gathering of human beings of common origin, history, usually with a common spoken language, most of whom are concentrated in a single country."
10. Rabbi Eliyahu of Vilna, *Devar Eliyahu: Commentary to the Book of Job*, Vilna, 1875.
11. With the following limitation: the expression "prefix X" must be between 5-8 letters. See [1]. This limitation was established in the first preprint of the paper: D. Witztum, E. Rips and Y. Rosenberg, "Equidistant Letter Sequences in the Book of Genesis", 1986. A later version of this paper was published in *Statistical Science*, v. 9, no. 3, 429-438.
12. D. Witztum and Y. Beremiz, *The "Famous Rabbis" Sample: A New Measurement*, May 10, 1998. A later version of this article (Sept. 1998) can be found at http://www.torahcode.co.il/english/pdf_files/new2e.pdf.
13. See [3], ch. 5 and 7.
14. See [3], Section 2.1.
15. Rabbi David Kimchi on Genesis 10:13, based on a *midrash* of the Sages. Some nontraditional sources also understand the name this way. See, for instance, *Encyclopedia Biblica*, v. 4, at the entry לוד, p. 439: "Ludim was one of the sons of Mizraim."
16. Examples similar to this one include: עם כתיים (Ramban, Numbers 24:21), עם פלשתים (Ibn Ezra, Zechariah 9:7). [See also עם היבוסים, *Metzudath David*, II Samuel 5:6].
17. Rabbi Samson Raphael Hirsch, *The Pentateuch*, Judaica Press, Ltd., Gateshead, 1982.
18. Rabbi Hirsch explains the terms לשון and שפה in his comments to Genesis 10:5, where the term לשון appears, and to 11:1, where שפה appears. According to

Rabbi Hirsch, Chapter 10 describes the origin of separate nationalities, a process that began with the development of “different dialects” of the single language. This process was only completed many years later, as described in Chapter 11, with the formation of completely separate languages. He comments further that while it is correct to speak of the German שפה and the French שפה, it is not correct to speak of the German לשון and the French לשון – “שפה would designate the language, French, German, etc., but לשון the way of pronouncing, the dialect.”

19. For example, a standard dictionary like the *New Dictionary* distinguishes five principle levels (see the Introduction to the *New Dictionary*).
20. The Babylonian Talmud, *Tractate Chullin* (137b), *Avodah Zarah* (58b). This principle is used by classic commentators such as Ibn Ezra, who states explicitly: “And so have they said that the language of Scripture is one thing and the language of the Talmud is another” (Ecclesiastes 5:1, Esther 8:17). The commentaries on the Mishnah also employ this rule [see, for example, *Tosfoth Yom Tov* on *Pesachim*, Ch. 1, Mishnah 2, *Temurah*, Ch. 5, Mishnah 1, and elsewhere].
21. Concerning the descendants of Noah it is written: “And all the land was of a single language [שפה אחת]” (Genesis 11:1). Ibn Ezra explains: “The meaning of שפה אחת is a single לשון.” Similarly Rabbi David Kimchi writes: “שפה אחת – that is, a single לשון, etc.” This relationship is made particularly clear in the Midrash *Kavod Chofeh* (כבוד חופה):

“... because all of them spoke a single language [לשון], as it is written: ‘And all of the land was of a single language [שפה].’ So, too, in the future Israel will speak a single language [לשון], as it is written: ‘For then I will convert the nations to a purer language [שפה ברורה]....’”
22. Even Shoshan (editor), *A New Concordance of the Bible*, Kiryat Sefer, Jerusalem, 1981.
23. S. Z. Havlin, “Report,” Jerusalem, 1996. Available at: http://www.torahcode.co.il/english/pdf_files/havlin1e.pdf
24. *The Responsa Database*, CD no. 5, Bar Ilan University, Ramat Gan, 1997.
25. It is interesting to note that *Encyclopedia Biblica* informs us that: “use of the term גוי in singular construct form is extremely rare in Scripture” (at the entry גוי, v. 2, p. 457).
26. Some examples:
 - For the term גוי כרתים (Zephaniah 2:5), and the commentaries note that this is not the name of a nation (according to *Targum Yonathan* כרתים is not a name at all, and it seems from the Vilna Gaon’s commentary to I Chronicles 18 that he concurs. According to Rashi כרתים is the name of a country, and according to Rabbi David Kimchi and *Metzudath David* it is the name of a family).
 - For the term לשון כשדים (Daniel 1:4). [The name כשדים (Kasdim) is usually rendered “Chaldeans” in English]. *Metzudath David* (ibid. 2:4) and others explain that this is actually a reference to the Aramaic spoken in Babylon. In other words, לשון is being used here only in the sense of “dialect.” In fact, the prophet Isaiah says of the Chaldeans, “This nation never was” (23:13), which Rashi explains to mean, “They were never fit to be a nation.” [In his commentary on the Torah (Deuteronomy 32:21) Rashi comments that the expression “not a nation” refers to the Chaldeans, who are “a nation without a name.”]
27. See [3], Section 2.2.
28. II Kings 9:20.
29. For example, the classic commentary *Metzudath David*, *ad loc.*

30. D. Witztum, E. Rips & Y. Rosenberg, *Equidistant Letter Sequences in the Book of Genesis*, Stat. Science, Vol 9 ('94), No 3, 429-438.
Available at http://www.torahcode.co.il/pdf_files/pub/wrr.pdf.