

Is there Hidden Information in Biblical Hebrew

Prof. Haim Shore

August, 2008

Copyright © 2008 by Haim Shore

1. Introduction

- “Thus was God observing the Torah and creating the universe” (Midrash Rabah, Bereshit, Parashah 1)
- “Bezaleel knew how to assemble letters with which heaven and Earth had been created” (Berachot, 55a)
- “It is well known from Kabbalistic literature that the letters of the Hebrew alphabet were created first. Thereafter, by using these letters, God created all the worlds, for it is written “In the beginning God created “Et” (in Hebrew, Et is a word combining first and last letter of the Hebrew alphabet; Or Torah, by Dov Ber, Hamagid Memezritch).

Reading Jewish sources, one is repeatedly confronted with a surprising revelation: our ancestors believed that there was a strong linkage between the physical world, as known to us, and the words of the original biblical language (Hebrew). This conviction is expressed not merely by general assertions, as quoted earlier, but also in various efforts, scattered throughout various Jewish sources, to produce real information about the physical world by analysis of the structure and the numerical values of relevant words in the Hebrew language. We should clarify right away: we do not refer here to Gematria. The latter relates to a common belief, adopted by rabbis and Jewish bible interpreters alike, that if two Hebrew words share the same numerical values – probably

there is a certain “secret” that binds them together, and it is up to Gematria to expose this common relationship. However, reading various Jewish sources one realizes that there was an additional belief, separate and unrelated to Gematria, that words in the Hebrew language provide hidden information about their respective objects, often independently of other words. The quotes above are various expressions to this belief, and a thorough search of ancient Jewish sources reveals that Chazal (an acronym for the Hebrew "Chachameinu Zichronam Livracha", "our sages of blessed memory") related to this belief seriously and introduced examples for its validity in abundance.

We relate to two examples. The numerical value for the Hebrew “Heraion” (pregnancy) represents the duration of human pregnancy (271 days; given in Midrash Rabbah, Bereshit, 20). Another example for “scientific” deduction, the validity of which we could not establish but which, nevertheless, present another instance for how Chazal had related to the Hebrew language, is that of Rabbi Shmuel in Midrash Rabbah. He relates to “I will greatly multiply the pain of thy child bearing” (Gen., 2:16): “Harbeh arbeh itzvonech ve-heronech”. Rabbi Shmuel claims that since “Harbeh” is numerically equivalent to 212, an embryo surviving 212 days in the womb will probably survive the whole pregnancy (Midrash Rabbah, Bereshit 20).

A fundamental supposition of the quotes in the opening paragraph and of these examples is that there is in Hebrew an additional layer, hidden to us, that provides information exposable only by relating to the numerical value of the word and possibly to some of its other aspects. For many years I was utterly convinced that this belief is based on superstition, a sheer myth unlike those provided by other religions and investigated by researchers of human cultures.

Over time, further examples that I have been exposed to, or those that I have detected independently, started to shake my earlier confidence that Chazal’s belief was just that. Let us introduce two further examples. They do not constitute “scientific evidence” but are intriguing as standalones, and in combination with other examples they may trigger perhaps a more thorough investigation of the subject of this article. The first example is

similar to the “Heraion” example. The word year in Hebrew, “Shanah”, is numerically equivalent to 355, the average duration of the lunar (moon-based) Hebrew year. A second example is the word “Ozen” (ear in Hebrew). This word derives from the same philological root as the Hebrew for “balance”. However, the scientific revelation that the mechanism responsible for the body balance resides in the ear was made only towards the end of the nineteenth century.

As I conveyed earlier, these examples and others have attracted my attention for many years. Although I related to them as curiosities, with no scientific basis, over the years my confidence that these are just “cases of coincidences” began to be undermined. In the summer of 2006 I reached the conclusion that the number of instances that I had assembled over the years have reached a “critical mass”, which justified putting these examples to print. The result of this conclusion is a book, published in the USA in the winter of 2007, titled: “Coincidences in the Bible and in biblical Hebrew”*. The name selected for the title well reflects my own attitude towards many of the examples given in the book: Coincidences, most of which still lacking scientific basis to exclude them from the status of “coincidence”, yet with such abundance that perhaps deserves a more serious reference.

While writing the book, something else had taken place. My background in applied statistics and in statistical modeling encouraged me to implement a more thorough analysis that would establish, once and for all, whether numerical values of Hebrew words indeed store additional information, relevant to the apparent meaning of the word. I assumed that one can expose this information by a proper statistical analysis. These were not just some vague thoughts. I had a rather clear conception of what needs to be analyzed and what was the right statistical approach. It is needless to emphasize that once a properly applied statistical analysis submit statistically significant results, with significance levels commonly accepted in science, then those supposed “Coincidences” cease to be so and in fact a scientific basis is formed for Chazal’s belief that there is hidden information stored in Hebrew words. To my great surprise, I have

found out that **probably** there is hidden information that supplements the exposed information submitted by Hebrew words. The eventual outcome of the statistical analysis conducted was a body of about 20 separate analyses. These are presented in the book, accompanied by graphical illustrations that allow the reader who is not versed in statistics to grasp the implications of the analyses.

The purpose of this article is to detail, in as non-technical terms as feasible, the analysis approach taken, and then introduce some outstanding examples for the findings of some of these analyses. We do not aspire here to convince the reader that these findings (and others given in the book) in and by themselves comprise enough scientific evidence to validate Chazal's belief. However, we do intend to open up a scholarly debate, and let the readers, either scientifically trained or otherwise, decide for themselves whether the examples given here, and in the book, are sufficient to support Chazal's attitude regarding the Hebrew language, part and parcel of Jewish tradition.

2. The Approach

Let us assume that we have two sets of data. One set represents daily measurements of environment temperature, taken over a span of several days (we have N such measurements, one per day). The other set has the same number of observations (N numerical values); however it is unknown what these measurements represent. A certain scientist claims that the N values in the second set are temperature measurements taken during the same period as in the first set, but these measurements were taken using a different thermometer and possibly also using a different scale. How should we check the scientist's claim? The answer is simple. Suppose that the second set represents the same temperature measurements, however on a different scale, for example Fahrenheit instead of Celsius. Let us arrange the measurements in each set in an ascending order. We obtain N pairs of measurements. Let us now draw the values in a graph, where the vertical axis denotes values from the second set and the horizontal axis values from the first set. Altogether we will obtain N points (as the sample size in each set). If the points align themselves on a straight line, it is apparent that the

measurements in the second set (whose meaning is unknown) represent the same temperatures as measured by the first set, however on a different scale.

What is so important about the fact that the points are aligned on a straight line? Let us recall how we convert Celsius (C°) to Fahrenheit (F°):

$$F^\circ = 32 + 1.8 (C^\circ) ,$$

namely, an equation of a straight line. The converse is also generally true: transition from one scale to another means application of a linear transformation. Thus, if two sets of values, when plotted on a graph, exhibit a linear relationship, one can assume that the data in these two sets of values represent measurements of the same “entity”, though on different scales.

This conclusion can be applied to examine scientifically Chazal’s attitude towards the Hebrew language. In the examples that will be shown below, we will examine sets of Hebrew words, bound by a common physical trait. In other words, for each word in the set there is a measurable value of the physical trait, unique to this word. We will exhibit the data in a graph, where the horizontal axis shows the Hebrew word numerical value, and the vertical axis shows the respective value of the physical property. We will examine whether the set of points align themselves on a straight line. Since the relationship between the numerical values of Hebrew words and the respective values of the physical trait is not a mathematical one (unlike the example of converting from Celsius to Fahrenheit), we need to implement statistical analysis that would determine that the proximity of the points to a straight line is close enough so that we can deduce that the observed alignment is not random but represents a real linear relationship between the variable represented on the horizontal axis (numerical values of Hebrew words) and that on the vertical axis (the physical property). A statistical analysis that would establish existence of a linear relationship would lead us to the conclusion that both variables actually measure the same “entity”.

And if this phenomenon recurs, over and over again, in biblical Hebrew, then Chazal's strong conviction that there is an additional layer of information hidden in Hebrew words, this conviction probably can be scientifically corroborated. My book displays about twenty such analyses. Three will be expounded below.

3. Three examples

First example: "Day, Month, Year" ("Yom, Yerach, Shanah")

May the numerical values of these Hebrew words represent the periodicity affiliated to the words? ("periodicity" is the major physical property that differentiates between the "entities" that these words stand for.)

To check this, we have to express the periodicity (or frequency) by a common measurement unit. For example, if we chose "Cycles per year", then the periodicity of "Day" would be: $(29.53059 * 12) = 354.37$ (the lunar month, on which the Hebrew calendar is based, is on average 29.53059 days). The periodicity of "month" will be 12 and of "year" 1. If "Day" was selected as the measurement unit, then the frequency of "Year" will be $1/354.37$ per day. We have opted to pursue a unit commonly used in science and engineering, denoted Hertz (1 Hertz is one cycle per second). Note that in ancient times the hour was divided into 1080 parts while in modern days we divide the hour into 3600 seconds. With regard to the results obtained from the statistical analysis, the actual unit selected is inconsequential provided use of this unit is consistent throughout the statistical analysis.

Table 1 displays numerical values of the words in the set and the frequency of the "object" that each word stands for. Note that a numerical value can be represented by any system (for example, the decimal system), or be expressed as a power value. Thus, the number 7 can be represented in two modes: $7 = 10^{0.8451}$. The number 0.8451 is denoted "the log of 7 to the base of 10." In fact, when numerical values in a sample of observations span several orders of magnitude, it is customary to represent these

observations, for statistical modeling purposes, by their log values. This is implemented here with respect to the physical traits of the first two examples.

Figure 1 displays the three points whose values are displayed in Table 1. On the horizontal axis numerical values of the Hebrew words are registered (DNV values), and the vertical axis displays respective values of frequency, on a natural log scale (log scale to the basis of "e"). We realize that the points align themselves on a straight line. If the points were aligned on an exact mathematical line a correlation of -1 would have been obtained. In practice we have obtained a linear correlation of -0.9992 (!). The actual statistical significance obtained is 2.5%. In plain non-technical language one can say that the probability that the points would align themselves, the way they did, by chance alone is very small (only 2.5%). Generally speaking, obtaining significance levels smaller than 5% is accepted in science as evidence for lack of randomness (in a more scientific parlance, as evidence that under the hypothesis opposite to the claim that we are testing, namely, the hypothesis that results have originated in sheer randomness, the probability of the results actually obtained is only 2.5%).

From a purely scientific perspective, one may doubt that the alignment of the points near a straight line is a coincidence.

A further analysis of this example will be incorporated in a computer simulation experiment, conducted with regard to the first two examples. This experiment will be described following the second example. Note that not all analyses in the book are confined to three points (like in the first two examples). Some analyses are based on data sets of up to seven observations. The third and last example will demonstrate the phenomenon that we relate to for a sample set of five observations (N=5).

Table 1. Data for the first example (“Yom, Yerach, Shanah”). DNV represents the numerical value of the Hebrew word. E-05 means 10^{-5} .

Word	DNV (Date Numerical Value)	Frequency (No. of cycles per second)	Log-Frequency (natural log scale)
Day	56	1.1574E-05	-11.3667
Month	218	3.9194E-07	-14.7521
Year	355	3.2661E-08	-17.2371

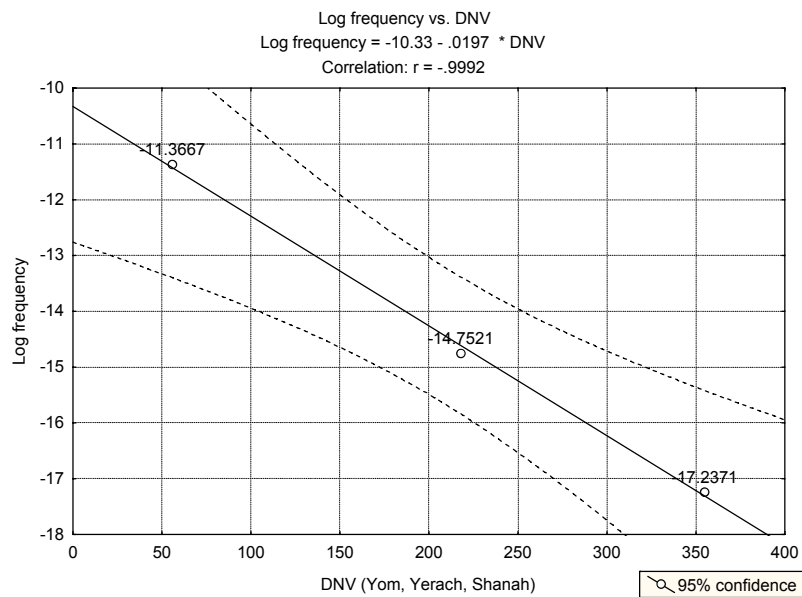


Figure 1. Periodicity (frequency, on a natural log scale) of “Yom, Yerach, Shanah” as function of numerical values of the Hebrew words (on the horizontal axis).

Second example: “Moon, Earth, Sun” (“Yareach, Eretz, Shemesh”)

In this example we examine Chazal’s conviction (that Hebrew words store information about our world) with regard to a central physical property that distinguishes between the three celestial objects in the set: their size, as given by their diameters. As in the earlier example, due to differences in orders of magnitude between the diameters of the three objects, the diameters will be registered in the plot on a log scale.

Table 2 introduces the data (diameters taken from NASA site), and Figure 2 displays the data, with ONV (Object Numerical Value) on the horizontal axis and the respective diameters, on a log scale, on the vertical axis. The phenomenon observed in the previous example is repeated: the three points align themselves on a straight line with a linear correlation of 0.999 (for an exact mathematical relationship a value of 1 would have been anticipated!). The actual significance level obtained is comparable to that of the earlier example (2.9% vs. 2.5% for the first example). Note that significance levels obtained for larger data sets are expected to be much smaller. Indeed for other analyses, given in the book (for example, with $N=7$) the statistical significance levels are appreciably smaller, thus lending more evidence to corroborate the overall conclusion that may be drawn from the totality of these analyses. The third example will relate to a sample size of $N=5$.

Table 2. Data for the second example (“Yareach, Eretz, Shemesh”). ONV represents the numerical values of the Hebrew words.

Word	ONV (Object Numerical Value)	Diameter (km)	Log-Diameter (natural log scale)
Moon	218	3474.8	8.153292
Earth	291	12756.28	9.453779
Sun	640	1 391 000	14.14553

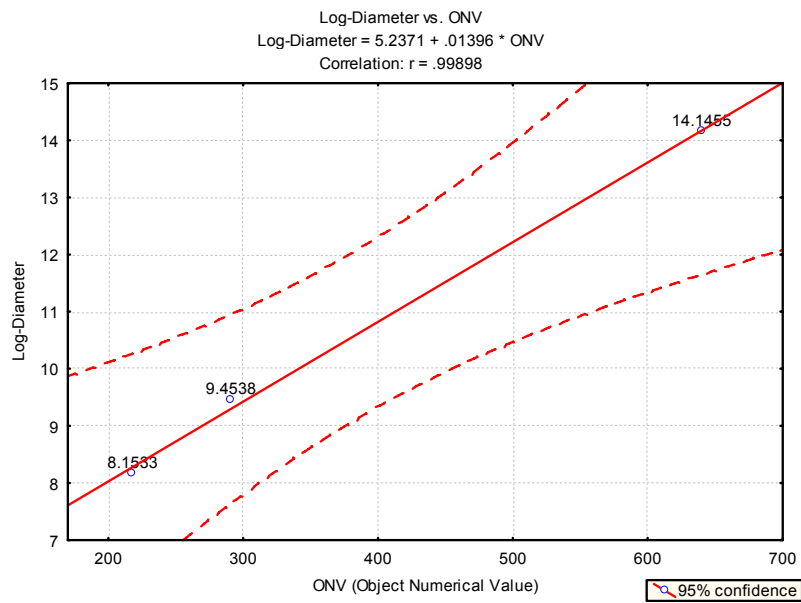


Figure 2. Diameters of “Moon, Earth, Sun” (on a natural log scale) as a function of ONV, the numerical values of the Hebrew words (on the horizontal axis).

A Computerized simulation experiment

The reader may justifiably wonder: may we obtain the same results if we had switched the two physical properties, namely, call “Moon” - “Day”, “Month” - “Earth”, “Year” - “Sun” (all referring to the Hebrew words). The reader may wish to carry out this experiment. We realize that the alignment of the points on an exact line vanishes, and the statistical analysis no longer delivers significant results.

To complicate things even further, we have planned the following computerized simulation experiment. Let us create sets of three Hebrew words each, where each word is composed of exactly three letters (as in the two examples above). However, the letters for each word will be randomly selected by the computer. We will repeat this experiment many times, and each time we calculate the correlation between the numerical values of the three words in the set and the respective values of the physical property.

Will the frequency of cases, where a correlation so high is obtained as in the two examples, be so large as to justify relating to the surprising results obtained for the two examples as some that could have happened by chance? We wrote a computer program (with software Mathematica®), and repeated the experiment a thousand times (namely, in each experiment there were a thousand repetitions). We have performed the experiment several times to ensure that the results obtained are indeed representative of the statistical population. We have also ensured that in “words” produced by the computer no letter will be repeated three times in a single word (we are aware of no three-letter word in Hebrew comprised of repetitions of a single letter). We have found that the proportion of cases, with the squared value of the correlation as high as in the two examples (or even higher), was in the range of 2-4%, namely very low probability.

Table 3 displays results from one experiment (with a 1000 repetitions), where the physical property relates to the second example (the log-diameters of the moon, Earth and the sun). The table introduces all cases where a correlation is obtained (either positive or negative) of size at least that obtained for the two examples. Altogether 1.7%

of the cases qualified according to this criterion. The most surprising finding from this table, however, is that the majority of “words” do not have any meaning in the Hebrew language, not even after permutation of the letters within the words (such permutations of course would not have altered the associated correlation). Furthermore, only one case in the table (case 5, one in a thousand with that high correlation) is associated with a trio of words that have at all some meaning in the Hebrew language.

Table 3. Trios of Hebrew “words”, produced by the computer to denote “Moon, Earth, Sun”, with correlations (either negative or positive) as large as those actually obtained (with the correct Hebrew words). Total number of sets produced during this experiment – 1000.

Case	“Moon” (value)	“Earth” (value)	“Sun” (value)	Correlation
1	(109) (ט, ג, ט)	(210) (ל, צ, צ)	(620) (ת, כ, ת)	0.9998
2	(490) (ל, ת, ס)	(409) (ר, ר, ט)	(137) (ל, ק, ז)	-0.9999
3	(36) (ח, כ, ח)	(84) (ד, ל, ו)	(251) (ר, א, ג)	0.9999
4	(200) (ג, ק, ו)	(161) (פ, פ, א)	(48) (א, מ, ז)	-0.9991
5	(51) (ל, א, כ)	(117) (ע, ז, מ)	(360) (ע, צ, ר)	0.9999
6	(16) (ז, ה, ד)	(90) (ס, י, כ)	(399) (ט, צ, ש)	0.9997
7	(404) (ק, ד, ש)	(330) (ק, ל, ר)	(16) (ד, ב, י)	-0.9996
8	(75) (ס, ט, ו)	(68) (ז, א, ס)	(48) (כ, כ, ס)	-0.9990
9	(97) (ט, ח, פ)	(175) (פ, צ, ה)	(471) (ע, ת, א)	0.9999
10	(49) (ל, ט, י)	(205) (ר, ב, ג)	(804) (ת, ת, ד)	0.9999
11	(14) (ב, ח, ד)	(58) (מ, ח, י)	(245) (מ, ה, ר)	0.9996
12	(53) (מ, ה, ח)	(180) (צ, מ, ג)	(602) (ב, ת, ר)	0.9999
13	(504) (ד, ת, ק)	(430) (ע, ס, ש)	(200) (צ, ג, ס)	-0.9996
14	(19) (ח, י, א)	(115) (ה, ג, ס)	(502) (ב, ר, ש)	0.9998
15	(66) (ז, ט, ג)	(167) (פ, פ, ז)	(600) (ק, ת, ק)	0.9996
16	(33) (כ, י, ג)	(30) (ט, א, כ)	(18) (ז, א, י)	-0.9998
17	(47) (ל, ז, י)	(100) (מ, כ, מ)	(295) (צ, ר, ה)	0.9999

Third example: Colors in Hebrew

May numerical values of Hebrew names for colors represent wave frequencies of these colors? Statistically testing this outrageous supposition serves here as a third example. This is a limited version of a much expanded analysis on the subject introduced in the book. For our experiment we have selected five colors that appear in the Bible, and for which there is little doubt among biblical scholars what they actually represent: red (“Adom”), yellow (“Tzahov”), green (“Yerakon”), blue (“Tchelet”) and purple or magenta (“Argaman”). These colors with their Hebrew representations are given in Table 4. A few comments about the Hebrew words are due. First, some names may be written with the Hebrew letter “Vav” or without. Inclusion (or exclusion) of this letter would not have changed the statistical results (introduced later). Final determination with regard to the “Vav” was based on frequency of appearance in the Bible (the most frequent form of the word was selected for the analysis). Secondly, the modern Hebrew word for green, “Yarok”, does not appear in the Bible. The only word close to represent green is “Yerakon” (Jer., 30:6). Most Jewish interpreters of the Bible (like Radak or Metzudat David) refer to this word as implying “green”. Note, that use of “green” to describe paleness, as in the above source from Jeremiah, is common in other languages, like English (where you say “his face turned green” to mean paleness). The modern meaning of “Argaman” is purple; however the close color magenta may qualify too. For reasons detailed in the book we have opted to assume that “Argaman” is magenta. Again, selecting the other interpretation would not have meaningfully altered the following statistical results.

Table 4 introduces the wave frequencies of the colors (in units of tera-Hertz, 10^{12} Hertz, where a Hertz is a cycle per second). Numerical values of the Hebrew words are also introduced. The first prominent observation that one gets from the table is that the Hebrew “Color numerical value” (CNV) and the wave frequency (WF) are arranged in the same ascending order. Since there are 120 possibilities to arrange five objects in any order, this implies that the probability of the two sets (values associated with WF and values associated with CNV) to align themselves by chance in the same order is $\frac{1}{120}$ or 0.83%. This is by far smaller than the critical value of 5%, given earlier, for achieving statistically significant results. An additional analysis, similar to the ones given for the previous two examples, is given in Figure 3 (one of several introduced in the book). The significance level achieved is 1.8%. Smaller significance levels are obtained for other analyses (not given here). For example, repeating the analysis of Figure 3 without “Red”, gives, for $N=4$, a correlation of 0.994 with significance level of 0.23%.

4. Conclusion

It is a tradition passed on from Chazal that words in Hebrew, which appear in the Bible, occasionally convey hidden information on top of the revealed information. We have related some examples for this tradition from ancient Jewish sources.

In this article, we have attempted, in as plain and non-technical terms as possible, to describe some quantitative analyses that demonstrate that perhaps this phenomenon of representing a major physical property in the numerical values of Hebrew words, perhaps this phenomenon is more prevalent in biblical Hebrew than formerly conceived. The variety of subjects that succumb to this characterization, as detailed in the book, makes it nearly impossible to relate to it as a curiosity devoid of any significance.

In the new edition of the book, published in 2008, a new chapter has been added that applies statistical analysis, based on principles not unlike those used here, to data given in the first chapter of Genesis. But this is a “story” that perhaps deserves to be a described in a separate composition.

*** Shore, Haim (2007, 2008). Coincidences in the Bible and in Biblical Hebrew. iUniverse, New-York.**

Comment: The research described in this paper has been conducted in the author’s private time, and it is not in any way associated with his research work at Ben-Gurion University of the Negev.

The personal homepage where this paper can be found (either in English or in Hebrew) is: <http://www.bgu.ac.il/~shor/index.htm>

® MATHEMATICA is a registered trademark of Wolfram Research, Inc.

Table 4. Data for the third example (“colors” in Hebrew). CNV represents the numerical values of the Hebrew words.

Word	CNV (Color Numerical Value)	Wave Frequency (WF)
אדום (Red)	51	443
צהב (Yellow)	97	520
ארגמן (Magenta)	295	546
ירקון (Green)	366	565
תכלת (Blue)	850	650

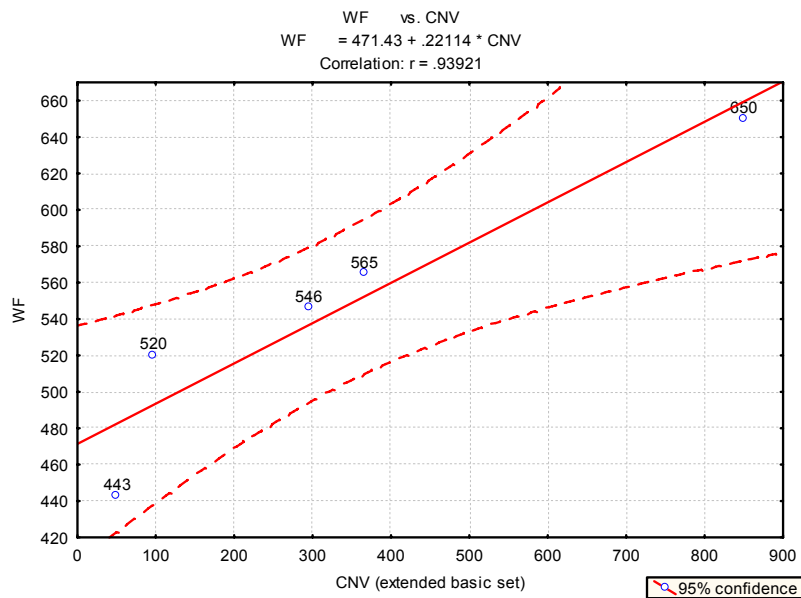


Figure 3. Wave frequencies (WF) of colors as function of CNV, numerical values of hebrew words 9on the horizontal axis).