

## Linguistic Connections among Torah Codes 29 November 2004

Art Levitt<sup>1</sup>, Robert M. Haralick<sup>2</sup>, Eliyahu Rips<sup>3</sup>

<sup>1</sup> Computer Research Analyst, Jerusalem, Israel

<sup>2</sup> Computer Science Department, City University of New York, U.S.A.

<sup>3</sup> Institute of Mathematics, Hebrew University, Jerusalem, Israel

artlevitt23@yahoo.com

### Abstract

This work presents a simple and verifiable method of testing the Torah code hypothesis. The hypothesis states that the Torah (the first five books of the Hebrew Bible) contains within it letter sequences (codes) that were created intentionally, as a form of communication to human beings, the intended receivers. Building on mounting evidence of multiple encodings for single events or concepts [1], our testing method measures the strength of linguistic connections between previous exceptionally noteworthy results. We first propose the general method. We then apply this method to Torah codes related to bin Laden and the Twin Tower attacks. The many significant connections that we find have an estimated probability of 1:500,000 of having occurred simply by chance. This strongly reinforces previous conclusions supporting the Torah code hypothesis.

### 1. Introduction

Torah code research is concerned with a particular type of letter sequence, formed by extracting equally spaced letters from a text. This is called an ELS (equidistant letter sequence).

ELS letter extraction is done by ignoring all punctuation and inter-word spaces. For example, the sequence "tin tops" can be found starting with the first "t" in the word "punctuation" in the preceding sentence, and using a skip distance of +4 (that is, counting forward every 4 letters from the starting position).

According to the Torah code hypothesis, logically or historically related words can be found as ELS's in the Torah, associated with each other significantly more often and in a more compact area than expected by chance. Several previous Torah code studies, pro and con, examined clustering of ELS's for famous rabbis and their dates and cities of birth or death [1-6], and many other historical events. More recent studies [7-8] focused on current events such as the Twin Towers attack (see the tables in figures 1-5; as is customary, the underlying text is arranged in a table of constant width; all codes are in the original Hebrew, with English translations provided).

One of the main challenges in this kind of research is to design experiments which implicitly demonstrate to an outside observer that the number of tests performed was very limited. For this reason, the current work first presents criteria for designating previous Torah code tables as foundational. These tables themselves are then used both as the source for deciding what to search, and as the areas of the text that are re-searched. We look for connections, such as the same words repeating within or between foundation tables. Under the null hypothesis of no Torah codes, we would expect that the connections found among the foundation tables would be no stronger than those found among collections of tables created from other texts.

### 2. Description of the general method

#### 2.1. Identifying Foundation Tables

Our first step is to collect previous foundation tables that all deal with a particular topic of interest. A code is accepted as foundational if it meets two criteria. First, its key elements must be in statistically close proximity. Second, it must be openly verifiable, as described below.

התאומים - כ"ג אלול

בראשית ברארלי מאת השם יסור אתה ארץ the twins תהו ובהו וחשך עלפני יהוה  
האדם בארץ וכלי צר מחשבת לבו רק רע כל היום י נחם ירו וכי עשה אתה אדם בארץ  
ירור אל אבד סלך למאד צומו ולדתו ומבי תאבין אלה ארץ אשר ארץ ואשך ואשך  
עלפני יסדו אברהם הלך עמם שלחם ירו אמרה מכסה אני ימאברהם אשר אני עשה  
פילגשו ושמרה אומה וילד גר אאת טבר ואת חסו ואת חשו ואת מעכהו יהי וחי  
ראלהן שמו the tower אלהן אבי massacre י צחק בנחלו ימצאו שם בארם  
יזכר אלרי סאתו חלו שמעאל י האלרי יסו ופתחתו חמהו תהו ותלבדו ותאמרו  
יעקב Ishmael ידבר חמו (9/11) elul י חשקה נפשו בבת כסותו ונא אתה  
נו ותסדר בגד י אלמנו ותהמע י הותכס בצעיף ותתעלף ותשבכתחני י יסאשר עלד  
ימת מבזאת בתחני פרעה אמצאו התי אסבו אחי כסה קטן הנהשלחו מכס  
ובני שמעון ימואלו ימיו וואהדו יכין horror לבן הכנעני יתו בני לו גרש  
ביו יחי יוסף מאה ועשר שני יסו יראו יוסף לאפריס בני שלש יסמבני מכיר בן מ  
לי עקב באלשדי ושמיי ירו קאנו דעתי להם גסה קמתי אבריתי אתם לתלה את  
יצא מעפר עהו יאמרו עבדי פרעה אלקי הודמתי היה זה לנו ולמו קשלת אתה אנש  
ומצרי סביאני who weiled הברכבו ובפרשי ווי סעמלאך האלרי חהלך  
ישמעאל עמד ברי עמו גסבר יאמי ונולסו יגד משה דברי העמלא ירו ויא  
About three thousand people fell יסאמרו אלה קי יסאמרו אלה קי יסאמרו אלה קי  
תקדש יסו היה על מצחו תמיד לצרו לקלסלפני ירו ושבתה כתבת שנתש עשית מצ  
רבו יעשו בני ילו יכדבר משהו יפלמן הסמבי ומהו אכשלש אלפא יסו יאמרם  
די סלקרשי צלעה משכן השני יתו חמשי יבר יחלקרשי המשכן לירכתי יסו יעש  
from the nation that day

Figure 1: “9/11” is the axis. Found on 1 October 2001 (“horror” and “who wailed” added November, 2004)

הריסות	36
ישאשריטמאולאי תחטאו <b>dead</b> נפש ההואמתורה	ד יט: כ
קהלכי את <b>ruins</b> ורטמאמי נדהלאזי קעלי וטמאה	ד יט: כ
ואוהי תהלהלחקתעו למזמזמי הנדהיבטבגדי	ד יט: כ
והנה עבמ הנדהי טמאעדעהיבוכלאשו <b>געבוטט</b>	ד יט: כב
מאיטמאוהנפשהנגעתתטמאעדערבו יבאובני יש	ד יט: כב
ראלכלעהמדברצן בחדשהראשו ון שבעעםבקדש	ד כ: א
ותמתשמר יםותקברש <b>shock</b> יהו יקהלו	ד כ: א
עלמשהו עלאהרן וירבעעםעמשהו אמרו לאמרול	ד כ: ב
וגוענו בגועאחי נולפני ירובבלמהבכאתתק	ד כ: ג
לירוראל <b>המדרזהלמ</b> ותשמ <b>mourning</b> וולמ	ד כ: ד
העליטנו ממצי <b>הב</b> אתנואלהמקוסהרעהזה	ד כ: ה
לאמק <b>Towers</b> <b>המפן</b> ורמו <b>Ishmael</b> שתותו	ד כ: ה
יבאמשוואוהן מפני הקהלאכלפתחאהלמועדו יפלו	ד כ: ו
עלפני הסו יראכבו ירוראלי הסו דברי רוראלמ	ד כ: ו
שהלאמר קאתהמטהוהקהלאאתהעד האתהואהר	ד כ: ז
ךודברתם <b>הסלע</b> ליעני הסו נתן מי ווהוצאתל	ד כ: ח
המסימון הסלעוהשקיתאתהעדו אתבעירו יקחמ	ד כ: ח
שהאת <b>horror</b> ירורכאשרצוהו ויקהלו משהו	ד כ: ט
הרן אתהקהל אלפני הסלעו יאמר להם שמעו נאהמרי	ד כ: ט
סמהן הסלעו הנוציאלכם <b>the Twin</b> דווי	ד כ: י
ךאתהסלע במטהו פעמיסו יצאומי סרביסותשתעד	ד כ: יא
הובעירו יאמר ירוראלמשהוואלאהרן יען לאהאמ	ד כ: יא
נתמבי להקדישני לעיני בני ישראל לכן לאתביאו	ד כ: יב

Figure 2: Based on Maariv headlines; “Ishmael” is new, a connector to figure 1. “Airplane”, “attack” not shown.

עון, פשע בן לאדן	2168
תנחושלשחושו ירוי וואחחאלהחרהמהורהלהחיהל	א ז: יג
י קוהקמתי <b>the city and the tower</b> ורעודממ	א ט: י
הוהלראתאתהעיר ואתהמגדל אשרבנו בני האדם	א יא: ה
איכי אשתך וואלמהאמרתאתהי הוא ואתהאקאתהלילא	א יב: יח
לאשרלך לאתמאראני העשרתי אתאברסבלעד ירק	א יד: כג
ואת <b>Wailing</b> <b>הוא</b> <b>smoke rising from the</b> וולזר	א יז: ז
מרשעאולי יש <b>earth as from a kiln</b> ארפתספה	א יח: כג
ככרו יראו הנהעלהקי טרהארקקי טרהכבשו ויה	א יט: כח
לגויאשימנו וכי זרערהו או ישכאמברהסב בקרוי	א כא: יג
<b>Shocked</b> אלנער	א כב: כב
<b>silence</b> ומאזבת	א כד: א
נט חואתרבקאחת	א כד: ג
שרנשבעתילאם	א כז: ג
קאולי ימשיא	א כז: יא
בברם <b>bin Laden</b>	א כח: א
מדיעודשבועשנ	א כט: כז
תחשפה <b>Crime</b>	א ל: ז
מצאאתהתרפים	א לא: ה
יעקבאחרינוכ	א לב: כ
לאשו <b>Sin</b> נובנ	א לד: יב
יעקבשניסעשר	א לה: לא
ותיעקבי וספק	א לז: ב
וירורו יאמרי הודהלאו נן באאלאשתאחי קויבם	א לח: ז

בן לאדן [-2168] מנא האלם ויללה [-2168]  
ממצא של בועז מצנר הי”ו

Figure 3: A phrase with bin Laden as anchor, found on 21 September 2001; original source for “wailing”, found in 3 other tables as “who wailed”, and “the tower” found in figure 1. “Sin” and “crime” are repeated in figure 5a. Inset: the “bin Laden-only”

portion of the table after expanding with half-width in November 2004, reveals “Ishmael” once again (see figures 1 and 2).

מי יללה? - ארה"ב יללה	36
מרולאמרולוגוענו בגועאחי נולפני ירורולמה	ד כ: ג
הבאתמאתקהל ירוראלהמדברהזהלמ ותשמאנחנו	ד כ: ד
בעירנו ולמה העליטנו ממצריסלהביאתנואלהמ	ד כ: ד
קוסהרעהזהלאמקוסזרעותאנהו <b>הפן</b> ורמו וומים	ד כ: ה
אין לשותו תויבאמשהואהרן מפני הקהלאכלפתחאהל	ד כ: ה
מועדו יפלו עלפני הסו יראכבו ירוראלי הסו יד	ד כ: ו
ברירוראל משהלאמר <b>Who wailed</b> והתעדהאת	ד כ: ז
הואהרן אחיך ודברתם <b>הסלע</b> ליעני הסו נתן מי	ד כ: ח
יוהוצאתהלמסימון הסלעוהשקיתאתהעדו אתב	ד כ: ח
עירו קמשהאמתהמלפני ירורכאשרצוהו וי	ד כ: ח
קהלומשהואהרן אתהק <b>Towers</b> <b>המפן</b> ורמו	ד כ: ט
מעו נאהמרי סמהן הסלעוהנוציאלכם מיסו ירסמ	ד כ: ט
שהאת ידו יךאתהסלע במטהו פעמיסו יצאומי סר	ד כ: יא
ימות <b>wailed</b> עירו יאמר ירוראלמשהוואלאהר	ד כ: יא
ן יען לאהאמנתם <b>The U. S.</b> עיני בני ישראל	ד כ: יב
כן לאתביאו אתהקהלהזהאלהארק אשר נתתי להם	ד כ: יב
המיריביה אשר רוב בני ישראל את ירורו יקדשבו	ד כ: יג
ישלחמשהמלכ ים מקדשאלמלך אדום כהאמראחיך	ד כ: יד
ירוראל אתה ידע אתאכלהתלאהארש מצאתנו וירדואב	ד כ: יד

מי יללה [-2] מו ארה"ב יללה [-37] מו  
מנדיל [71] מזד התאומים [36] מו

Figure 2a: The same area as figure 2

בין לאדן רוצח אלפי איש	3805
השבבאמתחתי נו בתחלהאנחנו ובאי סלהתגלעלעלי	א מג: יח
תן להם יוסף <b>התעלפ</b> פן עהו יתן <b>The Twin</b> קל	א מה: כא
נצרכנו למעשנו נהו יקרבו ימי ישראל למותו יקרא	א מז: כח
כמוי כוחו גלנו אתו ויצאוהי לך לאמראביך	א נ: טו
אברהםאלי יצחקואלי יעקבשלחני אל	ב ג: ג
סבזרענטו יהושפט <b>סגדלים</b> וקבשעלילע	ב ו: ו
לחאתעמי הנוני משליחך <b>הבעבד</b> ויבטן ויבתיך	ב ח: יז
עסוהעתירו לירוראל היכסו <b>סרמעל</b> וקאתהמותה	ב י: יז
תכספו מלתהי וואזי אכלבו ותשובו שכירלאי אכלבו	ב יב: מד
המיכמה <b>נאדר</b> בקדשנו וראתהלחע <b>who wailed?</b>	ב טו: יא
כהן מדן חתן משהאכל אשרעשהאלהיחלמשהו לישר	ב יח: א
תעשו ואתיאלהיכספו <b>thousands of people</b> וואז	ב כ: פט

פגוע [2] מו בין לאדן [3805] 4מ מי יללה [-6] מז

Figure 4: top-cited words from Haaretz; and connectors “who wailed”, “murders” and “thousands of people”



Figure 5: The end of another bin Laden phrase



Figure 5a: Continuation of figure 5 using half width

There are at least four alternative ways to create openly verifiable codes:

- The keywords to be searched as ELS's come from an independent, unambiguous and well-known source.
- A single ELS phrase string, starting with a well-known keyword as the "anchor" [8], is semantically and/or grammatically cohesive.
- A reinforcing news event occurs after a code is discovered.
- The code may contain a kind of self-imposed constraint that greatly limits the number of plausible, equally surprising adjoining ELS's.

Codes meeting both criteria (close proximity and openly verifiable) are by definition very rare. We thereby greatly limit the number of foundation codes

that can be studied, which avoids the possibility of under-reporting the number of failed attempts.

2.2. Finding simple linguistic connections

After identifying the foundation tables, we look for "connectors" between them. A connector is a word that occurs more than once in the collected tables (ELS and/or underlying text appearances), as indicated in many of the figures. For current measurement purposes, we analyze only a subset of all possible connectors; we use selection criteria which are both limiting and very basic. We start with a list of keywords gathered mechanically from the ELS's and verses previously highlighted in the foundation tables. We accept only the more difficult-to-find words – those which are at least 5 letters long (including any fully acceptable alternative spellings, of which there is only one, for "terror attack"). We search each foundation table for each keyword from our list, to determine if the keyword is near the main axis (vertical theme) of the table. We limit further by rejecting all diagonal ELS's, thus isolating for study the more striking parallel, collinear and perpendicular formations – found to be significant in their own right in previous studies [9]. We do not give credit for the original find of a word – only for subsequent (connecting) finds.

2.3. Estimating p-levels for each connection

For each connection, we are interested in the probability that the ELS found in the table - in its observed horizontal or vertical configuration – is in such a close proximity to the table's main axis word. This probability is obtained by comparing the actual table to many random tables created in a Monte Carlo run that extracts them from modern and ancient Hebrew texts, where we expect no codes (even the Torah itself can be used for comparison since we do not expect randomly chosen tables generated from the Torah to contain connectors to our topic). In each comparison table, this run assigns a random axis location, searches for ELS's for the keyword, and determines the proximity of the closest ELS (with appropriate configuration) to the axis. The probability sought is the fraction of comparison tables with closer proximity than the original table. To measure proximity, we use the WRR [2] mu value, for the better of two table sizes (original and half-width). The half-width table size results in the axis word appearing as a column that skips alternate rows. Therefore we limit all vertical ELS's to skips that are equal to or double the table width. This follows the precedent seen in figure 5a, which was the motivation for the current study.

## 2.4. Estimating an overall p-level

With a reasonably large data size, the binomial distribution gives us an order of magnitude estimate of the overall p-level. The number of attempts, and the number of connections verified to have a threshold p-level, are the parameters for this calculation. The number of attempts is the number of items in the keyword list times the number of collected tables, times the number of configurations attempted (for example, 2, if we separately analyze the simple vertical and horizontal configurations).

## 3. Case study

### 3.1. The foundation tables

We apply the method described above to a set of former Torah code findings related to bin Laden and the Twin Towers (figures 1-5) [7-8]. These findings conform to the criteria as stated in section 2.1 above. Each is openly verifiable in one of the 4 ways listed there: figure 2 follows method (a) by using the headline nouns from *Maariv from 12 September, 2001*, while figure 4 uses top-cited words from *Haaretz*; these are 2 of the top 3 Hebrew newspapers in Israel. Figure 3 follows method (b) by using bin Laden as the anchor keyword, finding a phrase surrounding it (noteworthy because of the extremely close semantic connection in the Bible itself between the words “sin” and “crime”); and figure 5 is a continuation of a second ELS phrase string surrounding another occurrence of this anchor [8]. Figure 1 follows method (c). It was found on 1 October 2001, just a few weeks after 9/11. Its highlighted verse was puzzling, because it implied 3000 casualties, and estimates in the news at that point were closer to 6000. Corroboration for the code’s figure came many weeks later. Figure 2a follows method (d) by using the dialog mode of code searching, originated several years prior to 9/11. The dialog mode has a built-in question and answer structure. We see highly appropriate words filling in this self-reinforcing structure (“who wailed” and “the U.S. wailed”).

### 3.2. The significant linguistic connections

There are many significant connections which meet our constraints; several are not even shown in the tables. Four connectors meet a .001 p-level threshold. One is for the ELS “who wailed”, in figure 1 (its original appearance is in figure 2a, and it appears again in figure 4). We also have “horror” in figures 1 and 2; and “Ishmael” in figures 1 and 2 (appearing yet again

in table 3); and “thousands of people” in figure 4 (matching the verse in figure 1 that is literally “three thousands of people” in Hebrew). Based on a variety of types of comparison texts (including random tables extracted from the Torah itself), the p-levels for the first two successes consistently fall between  $p=e^{-3}$  and  $p=e^{-4}$ , and the final two are at least 1-2 orders of magnitude stronger.

### 3.3. The overall significance estimate

The keyword list built from and searched in the 5 unique foundation tables contains 13 entries. Per section 2.4, limiting the configurations sought to vertical and horizontal, the number of attempts is  $13 * 5 * 2 = 130$ . From these attempts, we have 4 successes at threshold level .001. Using the binomial distribution, this results in a p-level of  $1.0e^{-5}$ , or 1 in 100,000. If we extend our observations to include cases with a skip of 1 or -1, then “the tower” in figure 1 is a valid connector for its original appearance in figure 3 (and it meets the .001 threshold). This is a third configuration, beyond the simple vertical and horizontal, so the number of attempts increases to  $13 * 5 * 3 = 195$ . From these attempts, we have 5 successes at threshold level .001, yielding a p-level of  $1.84e^{-6}$ , approximately 1 in 500,000.

## 4. Discussion

It is often difficult to pin down a probability if strict constraints are not used. Figure 5a is such an example. It contains three potential connectors at the bottom. They meet the spatial constraints but not quite the linguistic constraints, since they use a verb form for “revenge” rather than the original noun form, and since there is a connecting “and” between “sin” and “crime”. These are rather severe constraints, but they avoid the introduction of elements that can not be properly enumerated in the probability calculations.

Another intriguing discovery that is excluded here is a cluster of words that could be added to figure 1: “wailing is heard”, and “in the U.S.”. It is excluded because it uses a diagonal ELS, and is not the precise wording that was previously found.

There is also an entire table which repeats the themes of airplane, attack, twice, and Twin Towers, using alternative spellings or synonyms [7]. Again it is excluded because it does not meet the strict criteria.

Even without any of these extra discoveries, the result is highly significant, over and above the significance of

the original tables (most estimated between  $p=e^{-3}$  and  $p=e^{-6}$ ).

## 5. Conclusion

Although we have not yet approached the point of understanding full meanings or interpretations, the current study strongly suggests that the Torah codes phenomenon is real.

We have reviewed a series of highly significant foundation codes, and we have observed reinforcing connections between them with probability 1:500,000 - revealing an even stronger structure than previously seen.

We see that the timing of these significant codes - their cross-validation or clarification over months or years – demonstrate a kind of “living” quality that reflects the paradox from ancient Torah commentaries that all is foreseen, and simultaneously we are granted free will.

## 6. Acknowledgements

We wish to thank Alex Rotenberg for developing and sharing his codes software; Boaz Metzger and Igor Pisetsky for contributing several of the foundation codes; Harold Gans for his valuable comments throughout the process; and Leib Schwartzman for his pioneering work with the dialog mode.

## 7. References

- [1] R. Haralick (2003); Torah Codes: Redundant Encoding; <http://www.torahcodes.net/redun.html>
- [2] D. Witztum, E. Rips, Y. Rosenberg; Equidistant Letter Sequences in the Book of Genesis; *Statistical Science*, 9(3):429-438, 1994.
- [3] H. Gans (2001); Torah Codes Primer; <http://aish.com/seminars/discovery/Codes/codes.htm>
- [4] R. Haralick (2003); Testing the Torah Code Hypothesis; <http://www.torahcodes.net/hypoth.html>
- [5] D. Witztum; Torah Codes; <http://www.torahcodes.co.il>
- [6] B. McKay, D Bar-Natan, et al, *Solving the Bible Code Puzzle*, *Statistical Science*, 14(5):150-173, May 1999.
- [7] A. Levitt (2004); <http://www.torahcodes.net/alltwin.html>
- [8] A. Levitt, N. Bombach, et al (2004); Long Phrases; 2nd Annual Speech and Language Conference; U of Belgrade
- [9] A. Levitt (2002); <http://www.torahcodes.net/w2d.html>