H. Van Dyke Parunak, *Structural Studies in Ezekiel.* Ph.D. Dissertation, Harvard University, Dept. of Near Eastern Languages and Civilizations, Sept. 1978.

APPENDIX

This thesis makes use of two original graphical techniques that require some explanation. Neither constitutes a new methodology. Both are attempts to visualize and discover relationships that have long been acknowledged as important in exegesis. They have the advantage of reducing fairly large bodies of data to conceptually manageable size. Thus they offer ways in which modern computational facilities may remove some of the drudgery from exegetical studies. Both these techniques display data in a manner to take best advantage of a scholar's intelligence and creativity. Neither can make any claims to replace such judgment. In fact, they may on occasion raise more questions than they solve.

1. Translation Concord Plots

1.1. The Problem

Often in biblical studies, one needs to determine the concord between a word in the language of one of the Bible versions, and the Hebrew word or root which that word usually translates. In any version, such information is indispensable to effective text-critical studies, in determining the text that lay before the translator. In the Greek traditions, this information is of great importance to the New Testament and patristic exegete as well. It can be used, with proper caution, to establish a bridge between the theological vocabularies of the Old and New Testaments, and thus to trace the history of doctrine and interpretation. Because of the particular importance of the Greek tradition, the detailed analyses of HR[1] have made it a central tool of biblical research ever since it was first published in the last century. We will confine our attention to the Greek tradition, but note that much of what we develop is applicable to other traditions as well.

It is almost never the case, of course, that one word in Greek corresponds completely to one word in Hebrew. We desire a graphic way to represent the correspondences among Greek and Hebrew words, so that we can estimate how complete a correspondence is, and perhaps develop a typology of correspondences that will advance the study of translation dynamics. In principle, all that we need has been collected in HR. But the bare columns of references and

---

[1]On the limitations of this work, and on other studies that may help to overcome them, see now E. Tov, "The Use of Concordances in the Reconstruction of the Vorlage of the LXX," CBQ 40 (1978) 29-36. His study articulates very well the questions which this methodology helps to answer.

code numbers do not lend themselves well to discovering patterns of correspondence. Can we do better?

Let us set two goals for ourselves. First, we wish to display which Greek word translates a given Hebrew word most frequently, which second most often, and so forth. (Of course, we may also look at the different Hebrew words that lie behind a given Greek word, but the basic concept is the same.) Second, we know that sometimes the amount of data available to us is not all that we could wish. We wish to estimate the degree of certainty or uncertainty that we can attach to these equivalences. We will develop the techniques of estimating concordance and the associated uncertainty in the context of an example.

1.2. Ranking Equivalences

In our structural study of Ezek 14:5, we encountered the Greek rendering enthymēma where MT's gillûl led us to expect eidōlon. The difference in meaning is so great that we desire to probe for possible textual corruption in one tradition or the other.

A first step is to see what enthymēma represents in Ezekiel.[2] From HR, we learn that enthymēma translates gillûl

---

[2]Because the Greek OT is not stylistically homogeneous, it is methodologically important to consider each book separately. Were further study to suggest that, as in Samuel and Kings, different translation conventions governed separate sections of the same book, we would want to draw our data only from homogeneous sections. On the other hand,

fifteen times, and *ʿălîlâ* only three times. It would, indeed, seem that *enthymēma* predominantly represents *gillûl*. Table 142 seems a helpful way to visualize the data.

```
                         enthymēma

                         0         5        10        15
gillûl                   ------------------------------X
ʿălîlâ                   ------X
```

Table 142

But the clash in meanings of the two Hebrew words forces us to take a closer look. After all, *gillûl* and *ʿălîlâ* do not occur with the same frequency in Ezekiel. We should not look only at the raw counts (How many times is each word translated by *enthymēma*?), but also at the percentages (Of all the times that *gillûl* is translated in Ezekiel, what percentage of times is it rendered by *enthymēma*?). The important number for *gillûl* is thus not its fifteen occurrences, but (since *gillûl* occurs thirty-eight times in Ezekiel) 15/38, or 39.5%. Table 143 displays

---

if we were studying the background of a NT word, we might want to consider the Greek tradition as a unit for its impact on the vocabulary of early Christians, at least as a first approximation. In our discussion here, we will consider Ezekiel as our universe of discourse. Thus "100% of the occurrences of a word" means all of its occurrences in Ezekiel; "word x occurs five times" means five times in Ezekiel, and so forth.

enthymēma

```
                 0.%                                   39.5%
gillûl           :------------------------------------X
ᶜll              :------------------------------X
```

Table 143

this information.[3] Now we see that, far from gillûl predominating over ᶜălîlâ as a precursor for enthymēma, the two are about equal with one another. Both are translated by enthymēma in 30 to 40 percent of their occurrences. We will see shortly that the difference between them is not statistically significant.

A logical next step is to ask: "For each of the Hebrew words that translates enthymēma, what does the total collection of Greek translations look like?" A survey of the index of HR yields the data that are presented in Table 144.[4] Experience suggests that Table 144 is a reasonable

---

[3]The easiest total count to use in figuring percentages is obtained by counting from a concordance such as Mandelkern or Lisowsky. But strictly, we should use counts derived from HR, since some occurrences of a word in Hebrew may not be present at all in the Greek tradition. The plan for analysis which we are proposing will eventually lead us to gather the total count from the Greek point of view, anyway.

[4]In addition to ᶜălîlâ, Ezekiel contains maᶜălal once, in 36:31, in a syntactic frame identical with one used commonly for ᶜălîlâ. Thus we have lumped the two words together in these plots.

^{c}ll, Raw Counts

```
              0 1 2 3 4
epitēdeuma    :-------X
enthymēma     :-----X
eidōlon       :-X
hamartia      :-X
```

^{c}ll, Percentages

```
             0.%
                                                         31.25%
epitēdeuma   :-------------------------------------X
enthymēma    :---------------------X
eidōlon      :-------X
hamartia     :--X
```

Table 144

picture for the distribution of a word and its translations. There is one main equivalent (epitēdeuma), and some semantically close items (enthymēma, hamartia). The contamination with the notion of idols, though, is not absent.

Table 145, the plots of raw counts and of percentages for gillûl, gives quite a different impression. The semantic misfits are not dwarfed by the semantic fits, as was the case with ^{c}ălîlâ. Gillûl has no one dominating Greek translation, and the group of words which translate it most frequently are not semantically homogeneous. One gets the impression, quite graphically, that the Greek translator does not understand gillûl. And since three of the four

```
                       gillûl, Raw Counts

                     0        5         10         15
dianoia/noēma        :-----X
eidōlon              :--------------------------X
enthymēma            :------------------------------X
epitēdeuma           :-------------X


                       gillûl, Percentages

                   0.%
                                                        95.00%
dianoia/noēma       :----------------------------------------X
eidōlon             :-----------------------------------X
enthymēma           :---------------------------------X
epitēdeuma          :---------------------X
```

Table 145

words which translate gillûl also translate the root ᶜll, it can also be surmised that ᶜll is the root with which gillûl is being confused. We will discuss other arguments that tend to support this conclusion in the body of the exposition.

1.3. Measuring Uncertainty

It is clear that percentages, more than raw counts, give a reasonable picture of what is going on. But we do not want to lose sight of the raw counts entirely. For instance, the percentage of times that dianoia/noēma translates gillûl (Table 145) is much less impressive when we realize that we are dealing with only three occurrences of a word, compared with fifteen for enthymēma. If enthymēma had been used one more time to translate some other word, its percentage of devotion to gillûl would be little affected. But if dianoia or dianoēma had one other occurrence somewhere else in Ezekiel, its value would drop by nearly 25%, moving it below its two nearest competitors.

What is wanted is some indication of how confident we are that the percentage given is correct. We do not mean that the arithmetic is not correct. But what if HR made an error in identifying one occurrence of a Greek word with a Hebrew one? Or what if deeper study were to show that the Greek tradition they followed is not representative of the Greek tradition as a whole? Or what if further manuscript discoveries should alter the correspondence? How much might the percentage change?

Statisticians have studied the uncertainty that attaches to numerical estimates of this sort. The uncertainty is found to depend upon two parameters. First, the larger the number of samples, the less uncertainty in the values computed from them. This is reflected in the difference between enthymēma and dianoia/noēma noted above. Second, percentages around 50% are much easier to change, and much less exact, than those near 0% or 100%. (It is much easier to move from 60% to 70% on a school examination, than to go from 90% to 100%.) These two principles may be combined into a measure of uncertainty called the standard deviation. Random binomial trials may be considered a model for our problem. In a large number of binomial trials, over 95% of all outcomes may be expected to fall within two standard deviations on either side of the mean.

We record the standard deviation of each measurement on our plots as shown in the examples in Table 146.

```
                             ʿll

               0.%                                     31.25%
epitēdeuma      :                       I---------------X----
enthymēma       :          I----------X----------I
eidōlon         I-------X-------I
hamartia        I--X---I

                           gillûl

               0.%                                     95.00%
dianoia/noēma   :                                  I----X----
eidōlon         :                                I--X--I
enthymēma       :                              I--X--I
epitēdeuma      :                I----X----I

                         enthymēma

               0.%                                     39.56%
gillûl          :                                I------X----
vbs/ʿll         :                  I--------------X----------
```

Table 146

The dashed areas on either side of the central percentage (marked with an "X") indicate one standard deviation on either side of the mean. What is the significance of this dashed region? The percentage relation between a given Greek and Hebrew word, indicated by "X", is only an estimate. The precise percentage is uncertain, for reasons mentioned above. There is approximately a 68% chance that the true percentage is within the dashed region (within one standard deviation of the estimated location at "X"). In other words, a sizable possibility remains that the real percentage might even be outside the dashed area. If we were to make the dashed area twice as broad, to include two standard deviations on either side of the estimate, we would have a chance of slightly better than 95% of bracketing the true percentage. If the area were three standard deviations wide, we would be 99.74% certain that we had trapped the true value.[5]

---

[5]These percentages are derived from the normal distribution, which is a close approximation to the binomial distribution (to which our data belong) only for large numbers. (Our standard deviations are computed exactly, according to the binomial distribution. It is only the percentage valuation of them as confidence limits that is approximated from the normal.) If we were formulating a more sophisticated mathematical argument, we would need more exact confidence limits, to take the small sample size into account. But our purpose is to get a rough feel for the data. For this purpose, the simple device of plotting standard deviations for normally distributed data, and understanding in general how they define uncertainty, must suffice.

Obviously, if the "true" location of "X" can be anywhere within the dashed region, we are not justified in considering two plots distinct if the "X" of one falls within the dashed region of the other. If one "X" is farther than one standard deviation, but closer than two, to another, then the chance that they are really distinct is only 68%. (This is just the chance that the "X" with which we are comparing the other actually is within the fences set up at one standard deviation.) We probably would not consider two X's distinct unless they were at least two standard deviations apart, when we would be 95% certain that they really are different. If we are very conservative, we may want to insist on a separation of three standard deviations.

One may define the following rules of thumb.

a. If one X is within the dashed region of another, there is no significant difference between the two as far as translation commitment to the source word is concerned. Thus, in Table 146, one is not justified in citing an instance of <u>enthymēma</u> as evidence for <u>gillûl</u> rather than <u>ᶜălîlâ</u>. (A glance back at Table 142 will show how far we have come in making such a claim!)

b. If the dashed regions overlap, but the X's do not, we are in the one to two standard deviation range. The 68% chance that they are really distinct will probably not persuade us. But we will certainly not want to bet on the 32% chance that they are the same. We will not argue either way from the data.

c. What if the regions do not overlap? If their separation is greater than the larger of the two standard deviations, we can argue that they are really distinct with a certainty of about 95%.[6]

One should note that these conclusions may legitimately be drawn only from pairwise comparisons of words within a plot. For instance, in Table 146, the forward plot for gillûl suggests that eidōlon is virtually identical with both dianoia/noēma and enthymēma in the degree to which it represents gillûl. One is tempted to conclude that enthymēma and dianoia/noēma are therefore virtually identical with each other. But the underlying mathematics simply do not permit such "second generation" conclusions.

---

[6] We will not pursue here the mathematical justification for using the larger of the two standard deviations. The basic theory is outlined in F. Mosteller and R. E. K. Rourke, Sturdy Statistics (Reading, MA: Addison-Wesley Publishing Co., 1973) 309. See Robert McGill, John W. Tukey, and Wayne A. Larsen, "Variations of Box Plots," The American Statistician 32 (1978) 16 on the sizes of notches in notched box plots for an application similar to ours.

More sophisticated multiple comparison techniques are available, but do not seem justified for our purposes here. (We hasten to add that dianoia/noēma and enthymēma are very close together, certainly much too close to permit the inverse conclusion that they are significantly different from one another.)

In sum, when two X's come within one another's dashed areas, we may argue that they are indistinguishable. When their shaded areas are more than one standard deviation apart, we may argue that they are truly distinct. Intermediate separations do not justify either conclusion. These data are simply inadequate to answer the question. (Other data may of course break the tie.)

1.4. A Note on Interpretation

The plots for Greek equivalents of one Hebrew word often differ in surprising ways from plots of Hebrew equivalents of one Greek word. Consider, for instance, the relationship between tôᶜēbâ and epitēdeuma in Ezekiel. Concord plots for both words are shown in Table 147.

In chapter 8, the question arises whether these words may correspond to one another. That is, given epitēdeuma in the Greek version where MT has tôᶜēbâ, does the Greek reading challenge the Hebrew, or not? If we were to consult

```
                    tô^cēbâ

                   0.%                                       61.71%
bdelygma            :                                   I-----X---
anomia              :                              I---X---I
epitēdeuma          :  I-----X-----I
hamartia            :I--X--I
asebia             I--X--I
anomos             I--X--I


                    epitēdeuma

                   0.%                                       36.76%
vbs/^cll            :                          I--------------X---
tô^cēbâ             :                    I------------------X------
gillûl              :                  I-----X-----I
```

Table 147

only the plot of the Greek equivalents of the Hebrew word, we would be inclined to insist that the Greek translator did not read tô^{c}ēbâ. Otherwise, it would seem, one of the more common Greek equivalents would have been used. But when we see the plot of the Hebrew sources that lie behind the single Greek word, our impression changes drastically. Tô^{c}ēbâ is fully as likely a translation source for epitēdeuma as is any other Hebrew word. Clearly, if we are to use these plots intelligently, we must ask about the difference between plots based on one Greek word, and those based on one Hebrew word.

The exact nature of this difference is a matter for further study, as is the whole system of translation dynamics which these plots display. But we may suggest some a priori theories of just what the plots represent.

In what situation is one Hebrew word meaningfully compared with a group of Greek words? The translator of a Hebrew text into Greek is faced with this question continually. Of the several Greek synonyms for a given Hebrew word, which is the most suitable? The answer may depend on various criteria. Contextual features may indicate which nuance should be recorded in the translation. The translator may strive to preserve the concordance of the text, as did Aquila. On the other hand, if there is a stylistic aversion to repetitions of the same word, as with the translators of the English King James Version, one may seek to find a synonym precisely for the purpose of avoiding repetition.

Thus, the plot of the Greek equivalents of one Hebrew word sketches for us the history of the preferences of the translator,[7] moving in the forward direction from Hebrew to

---

[7]The reference to "the translator" is an oversimplification. Even if we could be certain that we had one translator's work before us (and we do not even suspect that this is the case), we could not say to what extent its renderings are influenced by previous translation traditions. But in the absence of clear knowledge on the identity and contribution of various translators, it is not unreasonable or without value to adopt the model of a shared set of translation conventions, as represented by the preserved evidence. If subsequent study succeeds in separating translation styles, this should be taken into account in later concord plots.

Greek. For this reason, we may call it a forward plot. If a text under study presents us with the most common Greek equivalent for the expected Hebrew word, we may feel very comfortable in the textual security of the passage. But we should not be surprised to find rare translation equivalents used sometimes. We will be justified in examining what might motivate the translator, in such cases, to use a different equivalent than the usual one. We will not, on this evidence alone, be justified in disowning the Hebrew reading.

The plot of Hebrew equivalents for one Greek word is a much more artificial structure. It is difficult to imagine a setting in the dynamics of translation in which it could naturally arise. It is, rather, an almost unconscious product of the translator's activity. One may question whether the translator ever asks systematically, "What Hebrew words am I representing by this Greek term?" But if this plot does not depict the translator's thought processes, it does conform to those of the textual critic. The text critic is interested in moving in reverse from the data (a Greek text) to the underlying Hebrew source. A plot of the Hebrew equivalents of one Greek word, which may therefore be called a "reverse plot," is answering the question, "What might this Greek word represent?" Unlike the forward plot, which depicts the more or less conscious choices of the translator, the reverse plot is generated unconsciously, and thus is much more susceptible to statistical analysis. An unexpected Greek word in a forward plot may result from some translation canon which we do not understand, rather than from a faulty reading. An unexpected Hebrew word in a reverse plot is much stronger evidence that the Hebrew and Greek texts from which the data were gathered do not in fact represent the same tradition. For it is much more difficult to imagine any process by which the unusual correspondence could have been generated.

We consider it methodologically necessary, in examining a disputed reading, to draw up both forward and reverse plots. When they agree on the soundness or weakness of the proposed translation equivalence, we will be quite safe in following the conclusion to which they lead us. But when they differ from one another, the evidence of the reverse plot (moving from one Greek to many Hebrew words) is to be preferred over the evidence of the forward plot (which may be informed by translation conventions which we have not discerned).

1.5. Computational Methods

Here we describe, for the ambitious, the computational details necessary to implement this methodology. A Fortran IV program does everything described here, with the exception of the first part, "Gathering the Data." But those without access to a computer will find the accompanying description sufficient to verify our results and to analyze items not discussed in this dissertation.

1.5.1. Gathering the Data

It is helpful to visualize the data as a matrix whose columns each correspond to a Hebrew word, and whose rows each correspond to a Greek word. syntax! The number of times a given Hebrew word is translated by a given Greek word is recorded at the intersection of the appropriate row and column. One additional row and one additional column, labled "other," are included to record the total number of times the word represented by that row or column corresponds to other words that are not included in the matrix. The data matrix for our study of <u>gillûl</u> and <u>^{c}ălîlâ</u> is displayed in Table 148. Once these data are collected, if the computer program is available, the researcher need only enter them in the appropriate form, and specify the words

	ᶜll	gillûl	tôᶜēbâ	other
epitēdeuma	4	7	2	0
enthymēma	3	15	0	0
eidōlon	1	13	0	1
dianoia/noēma	0	3	0	0
other	1	0	2	0

Table 148

(either Hebrew or Greek) for which he wishes to have plots computed. The next section outlines the computational details.

1.5.2. Computing the Plot for a Word

As an illustration, let us pursue the computation of the correspondences for gillûl. We begin by copying out just that column, as indicated by column A of Table 149. (If we were studying a Greek word, we would copy out a row, but could turn it on its end to follow the format shown here.)

Alongside each Greek word, we note the total number of times that Greek word occurs in Ezekiel, in column B of Table 149. These totals are obtained by summing each of the rows in the original data matrix. Thus for epitēdeuma, we add across the top row of the matrix, 4+7+2+0 = 13. Note

	A	B	C	D	E
epitēdeuma	7	13	.54	.13	23+/-5
enthymēma	15	18	.83	.09	35+/-4
eidōlon	13	15	.86	.09	36+/-4
dianoia/noēma	3	3	.95	.12	40+/-5

Table 149

that the row totals are used when we are studying a single column, and the column totals are used when we are studying a single row.

Column C contains the proportion of each Greek word's use in Ezekiel that is devoted to gillûl. The most straightforward computation would be to divide column A by column B, giving the percentage of each Greek word that is devoted to the Hebrew word under study. But in computing the standard deviations for each of these percentages, zero cells in the matrix give unrealistic results, as we shall see. Thus we have "started" our counts, as advocated by Mosteller and Tukey,[8] and added 1/6 to all cell counts

[8] F. W. Mosteller and J. W. Tukey, Data Analysis and Regression: A Second Course in Statistics (Reading, MA: Addison-Wesley Publishing Company, 1977) 230; J. W. Tukey, Exploratory Data Analysis (Reading, MA: Addison-Wesley Publishing Company, 1977) 465.

(column A) and 1/3 to all sums (column B) before dividing. This makes little difference for most items. For <u>epitēdeuma</u>, column C is not 7/13 = .5385, but (7+1/6)/(13+1/3) = .5375. With this information, the reader can verify column C.

It is in computing column D, the standard deviations, that the start is most important. If we denote the value computed in column C as "p", and the total number of occurrences of a word as N, then the standard deviation is

$$\sqrt{(p)*(1-p)/N}\ .$$

Consider the behavior of this value for a word or group of words, such as <u>dianoia</u>/<u>noēma</u> in the present example, which occurs only once or a few times, and always as the translation of the same Hebrew word. The percentage commitment of such a translation to its source word, 100%, is misleading. Because the word occurs so few times, its total devotion to one source may be simply a coincidence. After all, if <u>dianoia</u>/<u>noēma</u> had occurred four times in Ezekiel instead of three, and if the fourth instance had been a translation of a different Hebrew source, the percentage devotion to the majority source would be only 75% instead of 100%. Clearly, this is the sort of situation where the computation of a region of uncertainty is especially important. Yet when p is 100%, the standard deviation becomes 0, and the impression given is that the 100% is virtually certain. The use of 1/6 and 1/3 to start individual counts and totals, respectively, avoids the mathematical singularity of a 100%. (In the case of <u>dianoia</u>/<u>noēma</u>, the percentage devotion becomes about 95%.) Thus a meaningful standard deviation may still be computed. Of course, the percentages no longer have value as absolute numbers, but it should be questioned how much value one would ever want to ascribe to them absolutely anyway. They are valuable only in comparison with one another, and this relationship remains unchanged by "starting."

Thus, with p computed from started values as in column C, the values in column D are computed from the formula

$$\sqrt{(p)*(1-p)/(N + 1/6)} \; .$$

As mentioned before, our percentages are of no value as absolute numbers. They help us only in comparing a series of words in one language which are translation equivalents for a single word in another language. If the numbers in columns C and D are plotted directly, some plots consisting entirely of low percentages will be very small, and it will be difficult to evaluate the separation of X's and shaded regions. For this reason, and to make the plots more uniform in appearance, we scale each plot (consisting of the graphs for several words in one language that correspond to one word in the other) so that the largest X is always at the same horizontal position on the page. If plots are to be made by hand on graph paper, the length to which the plots should be scaled will depend on the paper being used. In this thesis, a length of 40 spaces is used, to facilitate typing. If G is the value of the highest X in a given plot, and if the plots are to be scaled to 40 spaces, then the values in columns C and D are multiplied by the factor,

$$40/G \ .$$

In our case, G, the largest value in column C, is .95, so that this factor is 42.10. Column E shows the results of this scaling, with the number before the +/- representing the scaled value of column C, and the number after representing column D.

2. Word Density Spectra

2.1. The Problem

So long as the study of symmetrical structures in the Old Testament is limited to verse parallelism and the patterning of short paragraphs, the existence of correspondences is easy to describe and verify. The repetition of a word or root is immediately apparent. And if an appeal to synonyms or grammatical constructions is