THE DIACHRONY DEBATE: A TUTORIAL ON METHODS[1]

A. Dean Forbes

Department of Hebrew
University of the Free State
E-mail: adforbes@post.harvard.edu

(Received 02/02/2016; accepted 05/10/2016)

ABSTRACT

Traditional approaches to the linguistic dating of Biblical Hebrew (BH) have produced many innovative results. However, because of inattention to the disruptive effects of textual noise and to the overfitting of textual features to restricted texts, these results have exhibited limited generalisability. In recent years, there have been proposals to include additional parameters in analyses. Lately, a construct from innovation theory, the s-curve, has been informally taken up by a few BH diachrony analysts. Not surprisingly, initial results have been approximate and provisional due to the idealised assumptions made. Future work along these lines must provide for features that are non-monopolising, non-monotonic, and fluctuating. Concurrently, the methods and inferences associated with traditional analyses have been questioned. For example, Young, Rezetko and Ehrensvärd have asserted that attempts to date biblical writings linguistically are *ab initio* illegitimate. I disagree.

TRADITIONAL APPROACHES TO LINGUISTIC DATING

Text-transmission noise effects

Traditional approaches to the linguistic dating of BH have produced interesting, but not compelling, insights. Results have been weakened in part by the habitual practice of not taking account of transmission noise in the texts. Four sorts of noise are associated with text analysis: transmission noise (due to inadvertent or intentional alteration of texts during transmission), feature noise (due to inconsistent mark-up of text features), class noise (due to misclassification of text blocks), and unmodelled-parameter noise (due to omission of parameters from the analysis). On the first three sorts, see Forbes (2012:11–12). The fourth sort is introduced below in the subsection headed "Writer demographics and social dialect". Figure 1 provides a representation of a process by which the texts may have evolved. Some indeterminate number of

[1] Presented at the 2015 SBL meeting in Atlanta.

sources (src_A^1 ... src_A^N, $N \geq 1$) were composed, combined, redacted, and transmitted (through the "composition and redaction channel"), yielding "Ur-scroll$_A$" as of notional date$_A$ – a fuzzy "crystallisation date" of transition from the "Biblical Period" to the "Masoretic Era".[2] During most of the Masoretic Era, copyists were enjoined to reproduce the text as received.[3] After many copyings (through the "transmission channel"), our witness to the text, scroll$_A$, emerged. Other scrolls (scroll$_B$, scroll$_C$, etc.) underwent their own transformations.[4]

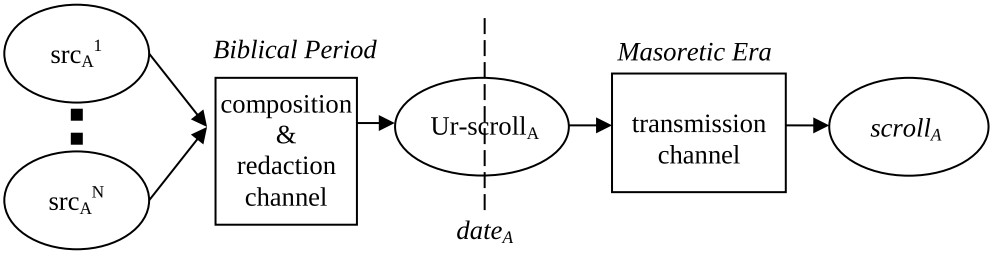

Figure 1. Possible evolution of a text portion[5]

Any reading in a scroll might result from noise, so generalisations based upon unique events are perforce unpersuasive.[6]

2 For insights regarding text transmission during the Biblical Period, see Carr (2011). On transmission as regards orthography, see Andersen & Forbes (1986:66–81).

3 Consider, for example, the words of Rabbi Ishmael (first century C.E.) in *Sotah* 20a: "My son, be careful, because your work is God's [Heaven's] work; because if you omit a single letter or add a single letter, you will be found to have destroyed the whole world, all of it!" Less strict views were also held. See Andersen & Forbes (1986:72–73).

4 For present purposes, this rendition of text evolution should suffice. For more formal analysis, models are based on information theoretic constructs (e.g., sources, sinks, Markov processes, etc.). See Andersen & Forbes (1986:105–125); also, Freedman, Forbes & Andersen (1992:111–124).

5 Refined from Forbes & Andersen (2012:129). Used with Eisenbrauns' permission.

6 Young, Rezetko and Ehrensvärd infer too much from too little when they argue that "the answer to the question whether there are any Persian loanwords in Early Biblical Hebrew texts can, on the basis of Deut. 33:2 alone, be answered with an unequivocal 'yes'" (2008/I:305).

Overfitting and non-generalisability

Another cause of weak outcomes in diachronic analysis involves overfitting of textual features to restricted texts, resulting in limited generalisability. Overfitting occurs when one adds more and more features the better to characterise a limited dataset. Such overfitted descriptions lack generalisability. That is, they fail properly to classify previously unseen datasets. Overfitting is a serious problem in pattern recognition.[7]

To gain an understanding of what overfitting and non-generalisability are about, a made-up example from clinical medicine may be helpful. Consider the initial diagnosis of systemic lupus erythematosus (SLE). The diagnosis was long based on the number of criteria (from a set of eleven) met by a patient. If four or more were met, then SLE was the diagnosis. This test was correct for 85% of patients having SLE and correct for 95% of non-SLE patients (Firestein et al. 2012:1284). Suppose that in an attempt to improve diagnostic accuracy, we apply the criteria to twenty SLE patients (too few!) and find that three are incorrectly diagnosed because they meet only three criteria. If two of the trio are redheads and the surname of the third ends with "-cilli", then we can add "IS-A redhead" and "surname ends with -cilli" to the criteria, and all twenty will be correctly diagnosed. But while these added criteria improve the diagnostic accuracy in this small group of ("teaching") SLE patients, they will be quite misleading when applied to other previously unseen ("testing") groups of SLE and non-SLE patients.

Overfitting features usually leads to classification rules lacking generalisability.

Improving accumulation

Central among innovations in diachronic analysis are Hurvitz's criteria determining the admissibility of evidence: (late biblical) distribution, linguistic opposition, extra-biblical sources, and accumulation (2006:194). According to Hurvitz's theory, for a given text-describing feature to be judged admissible, it must meet the first three criteria.

---

7 Abu-Mostafa et al. (2012:119) offer this insight regarding overfitting: "The ability to deal with overfitting is what separates professionals from amateurs in the field of learning from data."

Once admissible features are identified, they must be submitted to the test of accumulation, which specifies that “it is only on the basis of a heavy concentration of late linguistic elements that a late dating may be securely established for texts of unknown age” (Hurvitz 2000:153). Such simple accumulation tests, where the test statistic is a sum of unweighted binary variables, are very common in clinical medicine (Langbehn & Woolson 1997:2679).[8] Unweighted variables are rarely used in the physical sciences. Instead, a learned weight is assigned to each variable, taking into account its relative significance, and thereby approximating optimal classification results (“discriminant function analysis”).[9] Whether use of properly-weighted features is to be preferred over use of equal-weighted features (also known as unit-weighted features) depends on the goal of the analysis. If the aim is “to place an interpretation on the weights in the linear combination, [then] accuracy of the estimates of the weights and uniqueness of their definitions is important” (Hand 1997:154). Interpretive analysis “assumes that all the significant variables are in the model” (Kerby 2003:250) and that “sample sizes [are] larger than 100 [times the number of] predictors” (Dana & Dawes 2004:328). I seriously doubt that either of these requirements will ever be met in reliable BH diachrony studies.[10]

But all is not lost. Where predictive analysis[11] is concerned, the literature is encouraging. According to Wainer (1976:216), “when you are interested solely in prediction, it is a very rare situation that calls for regression weights which are unequal”. Or consider Cohen (1990:1306): “As a practical matter, most of the time, we are better off using unit weights”. One can rely on equal-weighted features, especially if the test of accumulation is appropriately modified. Consider these areas of possible modification:

- convert tallies to ratios or proportions to quantitate “concentration”;

---

8 The SLE diagnostic tally above is a sum of unweighted binary variables.

9 A probabilistic treatment of the “sum of weighted binary variables” problem may be found in Duda, Hart & Stork (2001:51–54).

10 As will be seen below in the subsection headed “Textual parameters”, there are ways of taking account of some independent parameters in the analysis, but inclusion of “all the significant variables” is beyond our present state of knowledge.

11 Put crudely, is a given text portion Early Biblical Hebrew (EBH) or Late Biblical Hebrew (LBH)?

- tally all instances of each feature; and
- make features as reliable as possible in preparation for classification.

Convert tallies to ratios or proportions to quantitate "concentration"

In defining accumulation, the "concentration" should be either an intra-corpus ratio of counts or a relative frequency, a count divided by a measure of corpus size.

Nearly fifty years ago Robert Polzin (1976) was remarkably perceptive in his handling of the biblical data. He often computed the ratios holding for pairs of alternants within corpora and then compared these ratios across corpora. When he dealt with single features, he computed densities – reckoned as counts per verse – and then compared these densities across corpora.[12]

Many later authors have settled for parade-of-examples exposition, omitting even counts. Others included counts but wrongly treated them as though they were comparable concentrations (Forbes 2012:17–18). Relative frequencies are rarely encountered. Commendably, Ian Young (2008:21) set about correcting this common lapse:

> Nowhere to my knowledge has an attempt been made to specify how … accumulation should be measured. ... In response to this problem I developed a simple test of accumulation. Plainly put, this counts how many different LBH features occur in a given stretch of text. Normally, this stretch of text will be of 500 words length, … so that samples are comparable.

So, rather than normalising counts by a measure of corpus size, Young prescribes a fixed minimal text block size, making the observed counts characteristic of these standardised minimal text blocks. Fair enough. Regrettably, for estimating feature frequencies, as well as accumulations, the chosen block size was too small.[13] This

12 See Polzin (1976:29, 30, 37, 60, 61, 81, 84, et cetera). Only once did Polzin fail to reckon densities prior to comparison (1976:45). Whether the normalising denominator should be a verse count or some other measure of corpus size remains open. For other candidate normalisers, see Forbes (2012:22).

13 While the arguments leading to this conclusion are extensive, they should be accessible to

problem is, however, fixable.

Tally all instances of each feature

Young (2008:21) also addressed the problem of repetitions:

> Within [each block] we count how many different LBH features there are. We do not count repetitions of the same feature. Once an author has demonstrated the possibility of using a particular LBH form, there is no reason it cannot be repeated as many times as opportunity presents itself. Thus, for example, the LBH order of substantive before numeral occurs seven times in Ezra 1:9–11 simply because it is a list.

There is a flaw in this argument. We read: “Once an author has demonstrated the possibility of using a particular LBH form …” But, of course, the texts that we study have passed through the transmission channel (Figure 1), and so any given run of text may be contaminated as a result of transmission noise quite disconnected from the author of the text. Consider, as an example, this verbless clause from 1 Sam 25:2b:

וְלוֹ צֹאן שְׁלֹשֶׁת־אֲלָפִים וְאֶלֶף עִזִּים

“And-to-him flock three of-thousands and thousand she-goats.”

<noun> <numeral$_{\text{pl}}$> *and* <numeral$_{\text{sg}}$> <noun>

The first conjunct is of the LBH form specified as feature G49 in Young, Rezetko & Ehrensvärd (2008/II: 171);[14] the second conjunct is the alternant of the first. It will not do to argue that one or the other of these conjuncts has toggled from its original alternant sequence due to transmission noise. We have no way of knowing the individual histories of features. But if we were to encounter, say, a dozen noun phrases in a text block each with the sequence specified by G49, then that would increase the likelihood that transmission noise did not account for the observed state of affairs.[15]

motivated readers. See Forbes (2012:23–27).

14 “Increase of postnominal syntax in which a numeral, which is usually in plural, is placed after noun.”

15 Using Figure 1 as a model for how texts might have evolved, I view changes occurring after date$_A$ as random. Before date$_A$, systematic and/or haphazard changes probably occurred (Forbes 2016). If the random damage inflicted after date$_A$ was too great, no information

Consequently, we hold that each instance of a proposed LBH feature should enter into the accumulation sum.

Young's statement appropriately raises another significant issue: what to do about lists. Repeated features in lists can unduly inflate incidence counts if each occurrence is tallied. Young "solves" this problem by deciding that only the first instance of a feature within a given corpus is to be tallied.[16] This might, in fact, be the best way to handle feature replication in lists. My preference, however, is to determine which proposed LBH features actually occur in lists and to assess their influences.

Make features as reliable as possible in preparation for classification

Young (2008:21) has also commented that:

> Nowhere to my knowledge has an attempt been made to specify how much of an accumulation is necessary for a text to be LBH.[17]

Along with this observation, Young included various presentations of "Table 1. LBH Features in BH Texts (Descending Order)",[18] a wide ranging, ordered list of accumulations. Although he noted that "Hurvitz considered the appearance of seven LBH linguistic items in 749 words of text enough of an accumulation to indicate a post-exilic date for the Prose Tale [of Job]",[19] Young did not propose a cut-off.

Following Young's lead, we might attempt to define an optimal cut-off level in the accumulation totals above which lie LBH texts and below which lie EBH texts. Alternatively, we might explore means of classification more sophisticated than this.

Before endeavouring to do either, however, we should consider three sources of artefactual variation whose existence might subvert our classificatory efforts: too-

---

from earlier epochs would have survived. In the case of orthography, the channel has not randomised away all of the old information.

[16] Note that elsewhere five or more instances of a feature are required with no instance of the EBH alternant as one criterion for declaring that a "preference for" feature exists (Young 2008:22).

[17] Also, Young (2009a:622; 2009b:257) and Young, Rezetko & Ehrensvärd (2008/I:94; 2008/II:86).

[18] Young (2008:23–25; 2009a:623; 2009b:258); Young, Rezetko & Ehrensvärd (2008/I:132–36; 2008/II:87).

[19] Young (2009b:257); also Young (2009a:622).

small text blocks, ill-advised features leading to overfitting, and non-reproducibility of accumulation tallies.

Increase text block sizes

While Young's accumulation is a genuine advance, it fluctuates needlessly because it is based on too-small blocks of text. As I have shown elsewhere, doubled block sizes of ~1 000 words each should lead to much more stable results (Forbes 2012:25–27).

Winnow features to minimise overfitting

Features of the "cilli" sort conjured up in the SLE example above lead to overfitting. They must be banished from diachronic studies. Elsewhere, I proposed three criteria for judging features to be overfitted: "(1) they are not found in the biblical texts at all, (2) they have very limited distribution, or (3) their detection requires unacceptable subjective interpretation" (Forbes 2012:36). Extant or newly proposed features might be evaluated using these criteria.

Take steps to make results reproducible

Finally, there is the serious problem of lack of reproducibility. Young (2008:23) alerts readers that "although Rezetko, Ehrensvärd and I have thoroughly checked the samples, it is still possible that we have missed some forms in some of the samples, but the results are so clear that a slight adjustment here or there will not affect the picture that emerges". For my *Hebrew Studies* article (Forbes 2012), I computed a trio of accumulations to check selected Young, Rezetko & Ehrensvärd (2008) censuses. Table 1 below specifies the text blocks analysed and gives the accumulations reported by Young, Rezetko & Ehrensvärd (2008/I:136–137) in the column labelled "LDBT" and the quite different ones reckoned by me. I do not assert that mine are the true values as opposed to those found by Young, Rezetko & Ehrensvärd. But the results do suggest that accumulations merit careful vetting by dispassionate parties working with independently marked-up texts.

Table 1. Comparison of accumulations for three text blocks

Text Block	LDBT accum.	ADF accum.
1 Kings 22:6–35	8	13
2 Chron 18:5–34	7	13
Zech 1:1–3:1a	3	8

Accumulation or something better?

Early on, Hurvitz (1972:76) doubted that a cut-off for accumulation could be estimated: "Accumulation is relative. It is very doubtful whether we can mechanically apply statistical criteria to linguistic issues like these." Given how accumulation is created to this day, his statement remains realistic. But, once Young's modified accumulation is adopted and corrective measures are taken to deal with too-small text blocks, ill-advised features, and non-reproducibility of accumulation tallies, we should be able to estimate an accumulation cut-off. One way to do this would involve supervised learning (Hastie et al. 2009:486–487) and a formalism called "Classification and Regression Trees" or CART (Hastie et al. 2009:305–317).[20] But, as we shall presently see, "textual parameters" other than time muddle diachronic analyses that treat time as the sole parameter of importance. Consequently, I prefer to turn now to an examination of the effects of textual parameters on our texts, leaving CART analysis for another occasion.

TEXTUAL PARAMETERS AND BIBLICAL TEXTS

Some recent proposals for strengthening and/or illuminating BH diachrony studies suggest that analyses should include both temporal and non-temporal parameters.

Textual parameters

A census of textual parameters

The textual parameters approach is useful for investigating phenomena that affect text structure. According to Herring et al. (2009:1–3), it has these four characteristics:

- "[Textual parameters are] properties of texts and their contexts that condition variation within individual languages".
- "[They] are defined inductively, on the basis of systematic empirical study of older languages".
- "Because language structure can vary according to one or more textual

[20] The classic work is Breiman et al. (1984). See also Kerby (2003:252–257).

parameters, language itself is revealed to be a heterogeneous affair, and the notion of language evolution as a single process is thereby called into question".

- "Only texts from certain genres tend to be preserved from any given historical period, making it difficult [to separate] the effects of diachrony from the effects of genre".

Herring et al. (2009:1, 5, 21, 25) provide a compendium of textual parameters. By splitting both date[21] and dialect,[22] removing "translationese" as irrelevant, and gathering parameters under three headings, I have formed Table 2, the focus of the section below headed "Comments on textual parameters".

Table 2. Textual parameters

Writer-Specific:	writer demographics	social dialect	regional dialect
Text-Specific:	composition date	text type	genre
	register	orality	poeticality
Transmissional:	scribal influence	text cultural status	crystallisation date ("date$_A$")

One might argue that some of these parameters overlap. Note that this collection of parameters quite appropriately lacks non-linguistic correlates of possible composition date "such as references to Greeks, re-use of other texts, ideology, or tradition-historical considerations" (Davies 2003:162). Note that, in any given text, the writer-specific parameters might be intentionally and successfully disguised.

Countering reticence inducers

Readers will note that several of the textual parameters catalogued in Table 2 make sense for literary texts but not for transcripts of speech. Further, readers will likely have noticed that two of the textual parameters involve dates. These parameters are included to counter the "close-to-speech" compulsion and "no-dating-here" emphasis in Rezetko and Young (2014).

The close-to-speech compulsion

Rezetko and Young (2014:23, 217, 245–46) repeatedly and favourably cite Edgar

---

21 "Time" has been split into "composition date" and "crystallisation date".
22 "Dialect" has been split into "social dialect" and "regional dialect".

Schneider's "four basic requirements for written texts to be useful for variationist analysis". Schneider's first requirement is that "texts should be as close to speech, and especially vernacular styles, as possible" (Schneider 2013:59, citing Montgomery). The cited passage begins:

> Scholars investigating early vernacular varieties of a language that has long had a standardised written form face … challenges. … They must first discover texts written in a colloquial style, reflecting speech as much as possible (Montgomery 1997:227).

Since we are not "investigating early vernacular varieties of a language", one wonders why we need to emulate those who are. Why not simply accept that we are dealing with literary texts, shed the "close-to-speech" compulsion of hard-core linguistics, and take seriously the important insights of stylistics?[23]

The no-dating-here emphasis

Throughout their book, Rezetko and Young remind readers of one major conclusion reached in their earlier book, namely that "the linguistic dating of biblical writings [is] unfeasible" (Rezetko and Young 2014:2, 249, 597). In Chapter 2, readers with any lingering "dating urge" are made to realise that adherents of historical linguistics simply have no interest in dating: "Prescription, in the sense of 'linguistic dating of texts,' is seldom if ever on the mind of the historical linguist" (Rezetko and Young 2014:15). Consider also this earlier comment from Young, Rezetko & Ehrensvärd (2008/I:61): "Historical linguistics, rather than the dating of texts, is much more commonly concerned with the relative dating of linguistic features, i.e. linguistic change, and the mechanisms of such change".[24] The second assertion oddly is true. According to Nevalainen and Raumolin-Brunberg (2003:56), "there is still a great deal of truth in Chen and Wang's argument … that 'one of the most neglected aspects of historical linguistics, which professes to be a study of language evolving across time, is the time element itself'". A clue as to why this should be appears in Crystal's

[23] Where BH and traditional stylistics are concerned, a good beginning is Kawashima (2013/III:643–650).

[24] I have elsewhere discussed gradient time ("absolute dating"), ordered time ("relative dating"), and adjacency-based time ("cluster-based dating"). See Forbes (2014).

definition of historical linguistics: “Historical linguistics uses the methods of the various schools of SYNCHRONIC linguistics (including SOCIOLINGUISTICS and PSYCHOLINGUISTICS, especially in considering the reasons for language change)” (Crystal 2008:229). Since both sociolinguistics and psycholinguistics involve primarily observational data, researchers know when the data are collected. There are no dating inferences needed. Because of the nature of experimental data, the statistical programs written for research in these areas until very recently had no provisions to represent change over time (Paolillo 2002:16). In the rare cases where time did need to be included among independent variables, special statistical packages were used (Paolillo 2002:15–16).

Interference associated with non-temporal textual parameters

Excluded textual parameters can lead to “fuzzy” results. Consider the Andersen-Forbes analysis of spelling in the Hebrew Bible. In our work on BH orthography, it was clear to us that we needed to carry out our analysis not solely in terms of text portion, but rather at least in terms of a trio of parameters: (1) vowel type, (2) stress level, and (3) text portion. Other variables were examined for inclusion and eliminated from the analysis as insignificant contributors to outcome (Freedman, Forbes & Andersen 1992:21–24). No attempts were made to include in our analyses any of the parameters in Table 2. Using fairly sophisticated computational procedures, it was possible to produce a so-called “horseshoe curve” along which text portions were ordered in terms of an underlying gradient variable, said gradient being interpreted by us as time (Forbes & Andersen 2012:141–144). The horseshoe, however, was not a skinny ribbon; it exhibited considerable scatter. Also, some parts of books were puzzlingly separated along the horseshoe. We observed that “the scatter and, perhaps, the separations are important indications that some other variable or variables are operative” (Forbes & Andersen 2012:144). This observation is consistent with a comment by Herring et al. referring to a diachronic analysis of *anon* by Brinton: “The case of *anon* … provides empirical evidence that diachronic change can operate independently in different text types, thereby problematizing any study of language change that does not control for text type" (Herring et al. 2001:13). The orthography analysis should be carried out with, at least, text types included. The basic mathematics previously used should still suffice.[25]

Comments on textual parameters

I next take up the textual parameters grouped in Table 2 above. Handling each of them involves choosing one of these options:

1. Tag the texts with parameter-specific feature(s) and include in the analysis.
2. If a parameter is concentrated in certain text blocks, consider omitting these blocks from the analysis.
3. Ignore the parameter, tolerating the resulting unmodelled-parameter noise.

Each option has its downside: options 1 and 2 exacerbate the sparse-data problem,[26] while taking option 3 may well contribute to undue solution "fuzziness".

Writer-specific parameters

The writer-specific trio of parameters surely affects BH texts, but the extent is all but impossible to know.

Writer demographics and social dialect

Rezetko and Young (2014:241) got it right:

> As for the writers, we know almost nothing about them. They are anonymous or unidentifiable, and although we might be able to arrive at somewhat educated conclusions about some of their demographics (age, gender, ethnicity, rank, social class and network, education, occupation, ideology, etc.), we know far too little about these to be able to explain the distribution of particular linguistic variables consistent with social correlates.

If any of these parameters accounts for a significant part of the variability observed in

[25] Technical comment: We would just be working with contingency tables in a four-dimensional space.

[26] The more parameters in a statistical model, the larger the samples need to be.

our data, then not including them in our analysis will mean that the variability unaccounted for after carrying out our analysis will be larger than would otherwise be the case. I characterise this increased residual variability as unmodelled-parameter noise.[27]

Regional dialect

If one posits regional dialect as a relevant textual parameter affecting the BH texts, one is by implication accepting the existence of such a dialect. On the basis of BH features that have congeners in texts written in neighbouring languages, Rendsburg (2013/I:338–341; see 2003:5–35) has sketched instances of a dialect that he calls Israelian Hebrew (IH) for passages identified as IH. The standard practice has been to take Option 3: Ignore dialect in analyses. Selecting Option 2 (omitting IH blocks of text) might reduce the dataset size by as much as a fourth.[28] If analysts conclude that dialect is an important source of unmodelled-parameter noise in preliminary results, they should assess the effects of implementing Option 1 on result stability.

Text-specific

In the linguistics literature the text-specific parameters exhibit sizable terminological overlap (Virtanen 1992:294–296). Biber and Conrad (2009:21) comment: "The terms *register*, *genre*, and *style* have been central to previous investigations of discourse, but

---

27 In the context of orthography, Andersen and Forbes have elsewhere examined this issue in great detail (Andersen & Forbes 1986:155–308). For present purposes, perhaps the BH orthography example stated non-technically will be helpful. We have a database consisting of all of the vowels in the MT, each labelled as to its spelling realisation (plene or defectivi) – the binary dependent variable. In addition, each vowel is tagged with four features (the problem independent variables/parameters): its vowel type, its location in the canon ("portion"), the stress that it is under (low, middle, high), and its verse parity (whether it is in an even-numbered or odd-numbered verse). We can devise a set of statistical models and vary their internal weights to estimate which model best accounts for the observed spelling choices. The preferred model (preferred, in a quantifiable sense) will leave some amount of the observed variability unaccounted for. Now come the crucial points for this discussion. When we omit, say, stress from the analysis, the residual variability is significantly increased, telling us that stress "matters." But, when we omit verse parity from the analysis, the residual variability does not change significantly, telling us that verse parity is (of course) irrelevant.

28 Rendsburg (2003:8–9) reckons about one-fourth of the Hebrew Bible is IH by chapter count.

they have been used in many different ways".

I prefer the demarcations proposed by Herring et al. (2001:10) wherein "genre" (akin to "text category") is an essentially open set, while "text type" is a closed set for a given language, usually with 4–6 members. "Registers, to a large extent, … conform to criteria related to the communication situation" (Virtanen 1992:294). While references to "written and spoken registers" are quite common, I prefer to refer to "written and spoken language", using "register" as per Crystal:

> register (n.) (2) In stylistics and sociolinguistics, the term refers to a variety of language defined according to its use in social situations, e.g. a register of scientific, religious, formal English" (2008:409; emphases levelled.).

Aside from the composition-date parameter, I consider all of the text-specific parameters to be candidates for inclusion in analyses (Option 1). However, our limited data likely will preclude inclusion of all of the parameters in a reliable analysis. Instead, analysts probably will have to pick one or two parameters to include.

Composition date

Given what is known about scribal practices during the biblical period of the history of the biblical texts,[29] hopes of recovering composition dates using linguistic information seem misguided. Recovery of relative "crystallisation dates" may be another matter. See major section headed "Crystallisation-date inference and seriation" below and Forbes (2016).

Text type

Text types should be included in future analyses. In our preparation for so doing, the text types for various languages will need to be examined with respect to BH, the text will need to be tagged with posited text types, and the analysis will need to be formulated to include text type.

Genre

If one accepts the proposition that the genre parameter involves an open set (as do I), tagging and including the genres would in all likelihood be an enormous task. I

[29] See Figure 1 above. See also Andersen & Forbes (1986:31–65) and its references.

suspect that the payback might not be enough to make the effort worthwhile.

Register

Biber and Conrad (2009:21) echo the sorry fact that "it is important to be aware that there is no general consensus concerning the use of *register* and related terms such as *genre* and *style*". The state of affairs concerning BH is evident when one discovers that Khan's massive encyclopaedia mentions "language registers" on four scant and scattered pages (Khan 2013/I:318; 2013/II:199, 201; 2013/III:582.). *Quid nunc*?

Orality

Herring et al. (2003:17–18) define and situate orality aptly:

> A fundamental problem for the study of languages of the past is the degree of orality of a given text, that is, the extent to which it reflects the properties of the spoken language at the time … While contemporary linguistics privileges spoken language as a more basic and authentic object of study than written language, historical linguists have traditionally been constrained to work exclusively with written texts.

Frank Polak (2012:724–25) has investigated ways of inferring discourse typology via syntactic-stylistic clausal analysis. He has described his procedures for building a "discourse profile" to measure clause complexity (Polak 2006).[30]

Poeticality

Watson (2013/III:151–154) has provided a brief, balanced survey of the agreed and contested characteristics of BH poetry. Although the central role of parallelism seems settled, debate continues regarding meter[31] and the sharpness of the prose-poetry divide. These issues may be bracketed for present purposes. The critical point is:

> Different genres may prefer different syntactic structures [and] can create profound – and generally under-appreciated – difficulties in the investigation of the synchronic and diachronic syntax of corpus languages. This is especially true when different chronological stages of a

---

[30] His discourse profile metric awaits refinement regarding nested clausal constituents and attachment preference ambiguity. See Forbes (2014:Table 4).

[31] Should one be a "foot finder", a "syllable summer", or neither?

> language are tested in different genres … [Thus,] poetry might prefer the 'flair' of greater variety in word order, while didactic prose might stick closer to unmarked order (Hock 2001:174).

Along similar lines, estimating text entropy, I found that "poetry appears to exhibit more freedom in the ordering of syntactic units than do other genres" (Forbes 1987:69).

One way to overcome these difficulties is to label text segments with their prose/poetry status and carry out analyses (Option 1 in section headed "Comments on textual parameters" above). This formulation assumes that the prose/poetry distinction is binary. If one decides that poetry-prose is not binary but rather lies along a continuum, then matters get more complicated.

Transmissional

The first two of the transmissional text parameters next taken up are included in the list provided by Herring et al. (2001:1), "crystallisation date" not being included. Although each parameter may influence the characteristics of a received text, rendering them quantitative to allow tagging is difficult in our case, if not impossible. Nonetheless, their possible effects have important implications as to which methods are appropriate to our situation. We consider each in its turn.

Scribal influence

This quote from a Kafkaesque short story by Borges provides an extreme caricature of scribes: "Scribes take a secret oath to omit, interpolate, vary" (Borges 1962:71).[32] At the other end of the notional "care gradient" are the members of a scribal guild taking to heart the previously quoted admonition of Rabbi Ishmael: "My son, be careful, because your work is God's [Heaven's] work; because if you omit a single letter or add a single letter, you will be found to have destroyed the whole world, all of it!"

[32] "Los escribas prestan juramento secreto de omitir, de interpolar, de variar". More realistically, under the heading "Bashing the scribes: language change as decay", Suzanne Fleischman (2001:47) criticises a too-well-established tendency among medieval philologists to "ascrib[e] the disconcerting variation and 'errors' of the manuscripts…to the negligence, defective linguistic knowledge, or mechanical copying methods of scribes".

(*Sotah* 20a). Tov (2012:184–85) depicts these antipodal scribal approaches thus:

> *Free approach*. When creating new copies, scribes altered the transmitted text, first as authors/editors-scribes, and later as copyists-scribes.… Scribal freedom comes to light in the insertion of many small changes (additions, omissions, content changes).… This freedom is also reflected in changes in orthography and morphology, while careless copying brought about many mistakes, erasures, and corrections. *Careful approach*. Carefully copied texts usually disallow changes in content, language, orthography, etc. The 𝔐-group (including many 𝔐-like texts found at Qumran…) is the major representative of this approach, although changes had been inserted also in this text at an earlier stage, and its Samuel text is often corrupt.[33]

Where does 4QSam[a] fit? Tov (2012:109) puts it near 𝔊 but also associates it with his group 4, "a cluster of non-aligned texts". In his second edition, he was more explicit: "4QSam[a] holds a special position ..., since it is closely related to the *Vorlage* of 𝔊, while reflecting independent features as well" (Tov 2001:116). In brief, Tov asserts:

1. The 𝔐-group texts were carefully – rather than freely – copied, although "its Samuel text is often corrupt."
2. 4QSam[a] is fairly close to the *Vorlage* of 𝔊, but it also contains significant "independent features as well."

These observations have methodological implications as regards the focal text in (Rezetko and Young 2014) and as regards "diachronic inversion", taken up in the major section headed "Crystallisation-date inference and seriation" below.

Text cultural status

Rabbi Ishmael's sentiment quoted earlier was paralleled by Josephus in *Contra Apion*, dated sometime after 94 C.E.:

> We have given practical proof of our reverence for our own Scriptures. For, although such long ages have now passed, no one has ventured either

[33] Tov (2012:24–26) defines 𝔐-group = Proto-Masoretic Texts and the Masoretic Text. Further, he defines 𝔐-like = Masoretic-like Texts from Qumran (Tov 2012:xix, 31–32).

> to add to, or to remove, or to alter a syllable; and it is an instinct with every Jew, from the day of his birth, to regard them as the decrees of God, to abide by them, and, if need be, cheerfully to die for them (Capps et al. 1926/I:179 and 181).

Feldman (1992/III:985–986) has given reasons that this declaration should not be taken too literally. Tov provides perspective: "At some point, a careful copying approach developed for the transmitted text, but it is hard to know when this process began (precision is not necessarily connected to the inception of a sacred or authoritative status, as shown by the transmission of the Torah….)" (Tov 2012:184). Adopting Tov's viewpoint, it appears that tagging the text portions with a cultural-status feature value is not feasible now or, probably, into the future.

Crystallisation date ("date$_A$")

In a linguistic variation analysis that includes time as an independent variable, the data are tagged with the known times of creation. For example, in a study of how language variation is affected by language acquisition, Berdan (1996:224–230) explicitly took account of the time variable.

Our situation is very different. We want to infer when the careful copying process began. For each text portion, we ask: What was its (absolute, relative, or cluster-based) crystallisation date?[34]

Summary

Table 3 lists the proposed textual-parameter options based on the foregoing discussion. Table 4 tells what each symbol in Table 3 means and gives my ordering of the tasks as regards research priority.

[34] On this terminology, see – in a long line of commentators – Tov (2012:21, 182)

Table 3. Proposed textual-parameter options

Writer-Specific:					
writer demographics	⃠	social dialect	⃠	regional dialect	3→1
Text-Specific:					
composition date	⃠	text type	**Add**	genre	?
register	?	orality	3→1	poeticality	3→1
Transmissional:					
scribal influence	?	text cultural status	?	crystallisation date	**Infer**

Table 4. Meaning of symbols (and research priorities)

Infer: Infer crystallisation dates	1st	**Add:** Include text types	2nd
3→1: Develop for inclusion	3rd	?: Identifiable parameter?	4th
⃠: Unrecoverable parameter	NA		

S-CURVES AND CONFIDENCE INTERVALS

Recall that an s-curve graphs the time-course of the value of a variable, increasing slowly at first, then accelerating in the mid-range, and finally slowing down as its value nears an upper limit. Figure 2 shows an idealised s-curve. Innovations diffuse through social systems. When “the rate of diffusion is treated solely as a function of interpersonal communication or social interaction between prior adopters and potential adopters in the social system” (Mahajan & Peterson 1985:17), then the profile for adoption is an s-curve. Of the grammatical features proposed in the BH diachrony literature gathered by Young, Rezetko and Ehrensvärd (2008/II:162–178), over two-thirds refer to a textual characteristic increasing (x53) or decreasing (x7). These are candidates for fitting by s-curves.

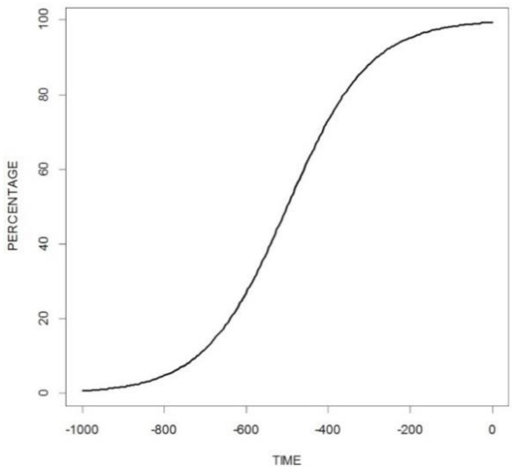

Figure 2. Idealised S-curve

In a paper given in Saint Petersburg in mid-2014, I identified three attributes of s-curves that have important methodological implications (Forbes 2016):

1. Non-monopolizing features.
2. Non-monotonic features.
3. Fluctuating features.

I next introduce these attributes and discuss their implications.

Possible non-monopolising features

The predicted change may not proceed to monopoly (Wang & Minett 2005:267–268). Then the s-curve upper limit will be less than 100%, and the theory will need a parameter specifying the upper limit.

Non-monopolistic behaviour occurs with stable alternation (Levin & Hovav 2005:186–205).[35] In English, for example, the dative shift occurred but has not (yet?) reached monopoly. We still have both “Jim gave the ball to Mary” and “Jim gave Mary the ball” (Levin 1993:29). Why one or the other form occurs in a given context has been analysed in detail (Bresnan & Hay 2008:245–259).

[35] On alternation in BH for the indirect object with אמר, see Andersen & Forbes (2012:344–345).

Non-monotonic features

Figure 3 shows Ellegård's data (1953:161) for affirmative declarative 'do' sentences. It is not an s-curve but increases and then decreases. Regarding this phenomenon, Vulanović concluded that: "The rise and fall of periphrastic *do* in affirmative declaratives can be viewed as two connected syntactic changes and their combination can be represented by what is essentially the same model. After all, the fall of periphrastic *do* in affirmative declarative sentences is the same process as the *rise* of the simple *do*-free constructions" (2007:124).

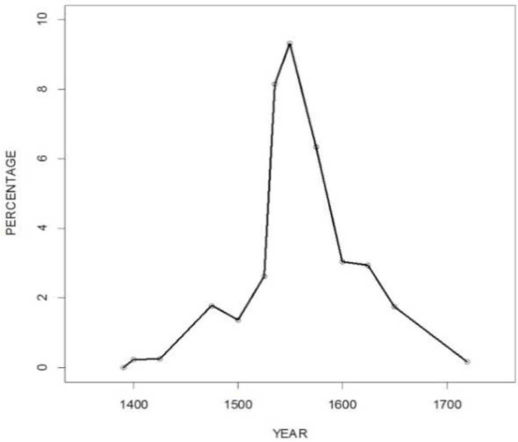

Figure 3. Affirmative declarative "do"

Vincent de Caën (2001:1–30) detected non-monotonicity in BH. He argued that final ה apocopation was later replaced by suffixation in some *wayyiqtol* verbs.[36] Holmstedt (2012:109–112) analysed the situation further.

With non-monotonicity, the inference from observed feature to inferred date, or date surrogate, is no longer one-to-one. It is one-to-many. Consequently, the "inverse solution" from dependent variable back to independent variable is not unique.

[36] De Caën noted that the application of one grammatical rule (apocopation) was supplanted, over time, by another (suffixation), leading to changes in terminal ה in certain *wayyiqtol* verbs.

Fluctuating features

As we will see below, features may fluctuate due to undersized samples or due to sudden corpus shifts. Either may cause very sizeable fluctuations. For example, consider Figure 4 reproduced from a classic paper by Altmann et al. (1983:110; used with permission). Here, many data points differ greatly from the fitted s-curve. As I have been unable to get the data underlying the figure, I do not know if undersized corpora or some other mechanism accounts for the fluctuations.

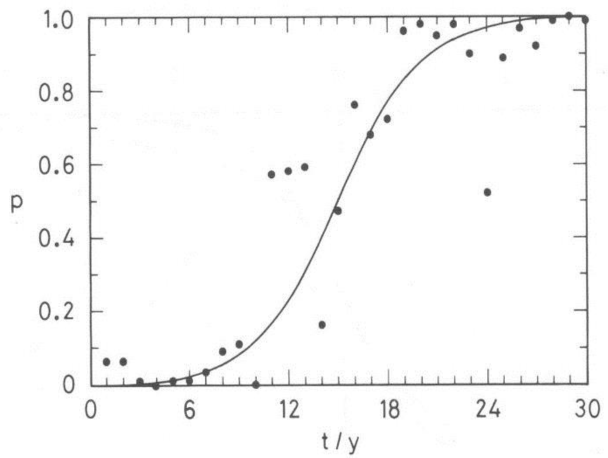

Figure 4. "Use of forms without ending"

The critical methodological point holds in any case: one may not simply order observations from smallest to largest and infer a temporal ordering, in effect seriating the data based on a single feature. Were we to do so for the data in Figure 4, the point at t/y = 24, for example, would be positioned far too early on any fitted s-curve.

Confidence intervals

The extent to which the fluctuations are due to undersized samples may be gauged by examining confidence intervals (henceforth, CIs).[37] Put loosely, CI theory allows one to compute an interval within which an actual population parameter lies, with some probability of being correct.

[37] For a BH-oriented introduction, see Andersen & Forbes 1986:8–14).

An example should help make concepts concrete. Suppose we need an estimate for the frequency of occurrence of the preposition -כּ ("like") in Jeremiah. For the sake of this example, we first divide the *Qere*-text of Jeremiah into 100-segment chunks, ignoring a final five-segment micro-chunk. For each chunk, we compute the relative frequency of -כּ and its associated 95%-CIs.[38] Figure 5 shows the results. The estimates (the solid black dots) scatter widely from the actual value. The CIs show that the actual value might lie anywhere between 0% and 10%. The sample standard deviation, a measure of scatter, is 0.92.

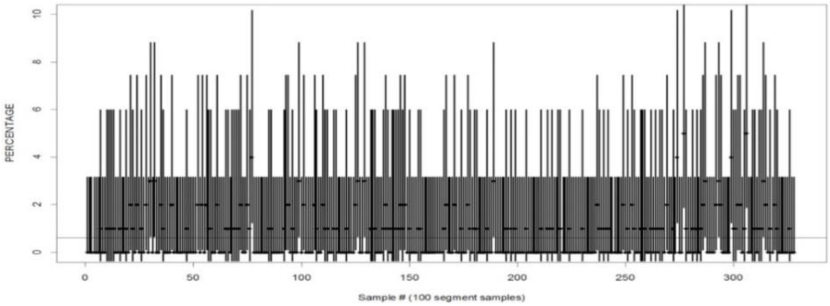

Figure 5. Relative frequency estimates for -כּ in Jeremiah (100-segment samples). The horizontal line at 0.622% is the population value. The vertical lines are the 95%-CIs.

Figure 6 shows the results for 400-segment chunks. Here, the true value might lie between 0% and 3%. The sample standard deviation is 0.44.

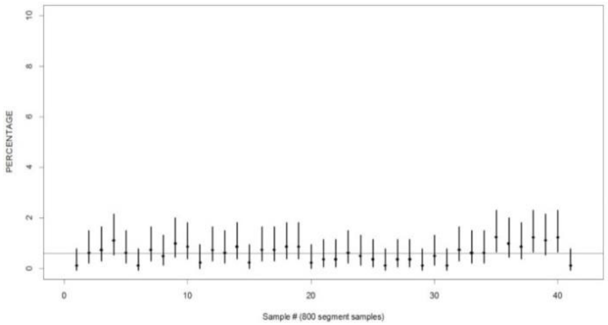

Figure 6. Relative frequency estimates for -כּ in Jeremiah (400-segment samples).

[38] Technical note: I use the Agresti-Coull interval as computed by the "propCI" function, part of the prevalence package in the R statistics suite. For practical advice, see Brown et al. (2001:101–133).

Figure 7 shows the results for 800-segment chunks. Here, the true value might lie between 0% and around 2%. The sample standard deviation is down to 0.34.

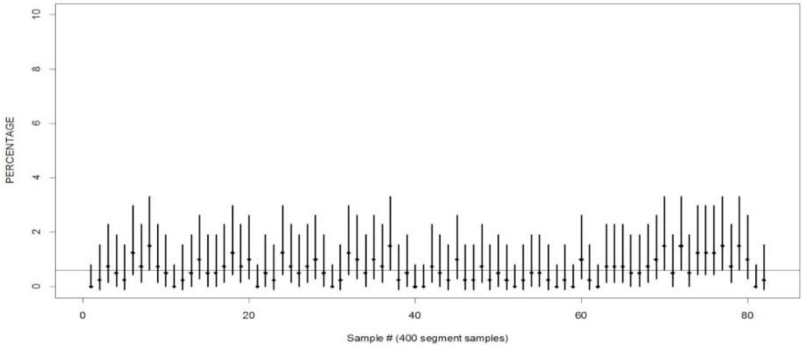

Figure 7. Relative frequency estimates for -כּ in Jeremiah (800-segment samples).

S-curve creation

That BH linguistic phenomena typically diffuse into use in rough accordance with an “s-curve” was a valid insight of Holmstedt (2012:118) and Cook (2012:91). However, the curse of fluctuating data muddled the details of their informal s-curves. To see what went wrong, consider Figure 8 showing 95%-CIs overlaid on Holmstedt’s nominaliser data (Holmstedt 2012:118; adapted with permission).

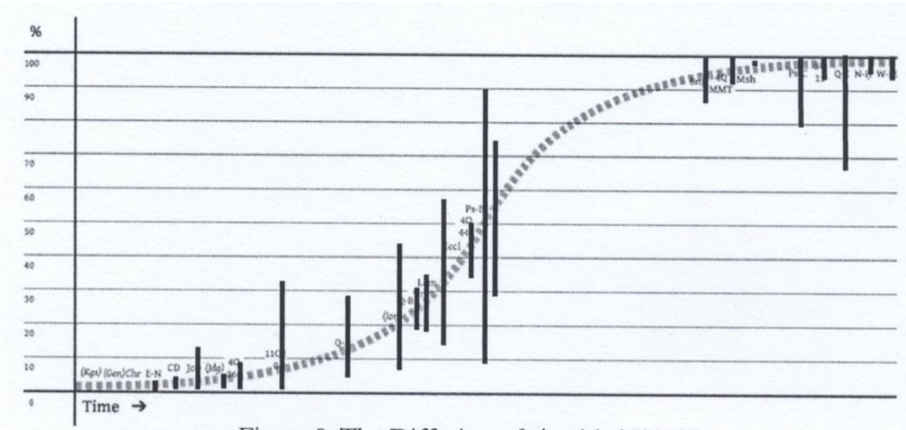

Figure 8. The Diffusion of שׁ, with 95% CIs

For the population out of which Ezra-Nehemiah might have been drawn (“E-N,” at lower left), the short CI indicates that the שׁ-proportion is probably close to zero. However, for the population out of which 4Q448 was drawn (midway up the curve), the actual proportion probably lies between 10% and 90%. A small sample yields a broad range of possible values.[39] Based on the CIs, many orderings of the corpora might be made. Ordering the texts from smallest-to-largest value is misleading.

REZETKO AND YOUNG: RELATIVE-FREQUENCY PLOTS

Rezetko and Young (2014:245–403) examine various linguistic features. They declare:

> *By selecting the 'late' verbs according to the results of previous scholarship, which largely decided that these verbs were late because they occurred in a predetermined set of 'late' writings, we have essentially engaged in the same sort of circular reasoning and stacked the deck against the 'late' writings* (Rezetko and Young 2014:318; italics in original).

This was done with "a main objective [being] to test the probability and categoricality of the conventional diachrony-oriented explanations for the variables. But time is only one dimension of variation. Other factors to consider are dialect, author/writer demographics … and so on" (Rezetko and Young 2014:319).

When Rezetko and Young plot the relative frequencies of ten verb lexemes on x-axes along which are ordered text corpora, the plots seem incapable of being connected so as to form monotonically-increasing curves, to say nothing of s-curves. Several mechanisms may account for the (wild) fluctuations of the relative frequency plots:

1. Reliance on small samples yields fluctuations and long CIs.
2. Introduction of non-biblical texts absent from the teaching data decreases generalisability.

[39] Each of the samples is quite small. All have fewer than 30 tokens, and 16 have fewer than 10.

3. Reliance on criteria having limited "diachrony community" support.

Let us briefly examine these factors.[40]

Tiny sample sizes

The BH diachrony literature is replete with linguistic features that occur very few times. Their relative frequencies of incidence therefore sprout long CIs. For the "late" verb lexemes studied in Chapter 8 of Rezetko and Young (2014), the central columns of Table 5[41] show the numbers of features occurring five times or fewer ("Sample-Size Counts") and the number of 100% relative frequency portions found for each lexeme ("100% Rel. Freq.").[42] We see from the "Sample-Size Counts" columns that around one in seven of the samples consists of a single token, with nearly one in two of the samples consisting of five tokens or fewer. The long CIs associated with the small samples allow vast numbers of "innovation curves" to be offered, as will be seen below.

Table 5. Some characteristics of ten selected verb lexemes

Lexeme	LDBT #	Sample-Size Counts					100% Rel. Freq.	Rel. Freq.		Support Level
		1	2	3	4	5		Ben Sira	DSS	
דחף/בהל	L40, L83	6	3	2	2	-	-	0%	13%	3, 2
בעת	L54	2	4	3	1	2	1	0%	4%	3
הלך	L93	2	3	5	1	1	-	23%	6%	2
זעק	L107	7	7	4	4	-	11	0%	100%	3
כנס	L153	5	-	2	4	3	-	0%	3%	4[a]
כעס	L154	9	4	4	7	1	1	0%	4%	2
עמד	L261	4	-	4	3	2	-	8%	8%	6[b]

[40] In the Rezetko and Young plots: a) the upper limit on the y-axis ranges from 15 to 225, making difficult both comparisons across plots and count estimation within plots; b) the number of text portions listed on the x-axis varies. For example, Figures 8.15 and 8.16 include 45 portions while 8.3 and 8.4 include 26.

[41] Here "LDBT#" refers to the lexeme designators in Young, Rezetko & Ehrensvärd (2008/II:162–214).

[42] In reckoning the "Support Level," all support from Qimron (1986) has been omitted because his comments are minimal, typically consisting solely of a reference to this or that publication.

קבל	L298	4	3	2	-	-	-	28%	9%	4
קום (*Piel*)	L302	5	5	3	2	1	1	0%	1%	5[c]
שׁלט	L349	9	7	-	3	3	2[d]	7%	1%	3
Column Totals:		53	36	29	27	13	16			

[a]- Hurvitz counted once in lexeme support level.
[b]- Comments by van Peursen are too peripheral to include his support.
[c]- The footnote in Sáenz-Badillos is too incidental to merit his inclusion among the supporters. The Joüon-Muraoka reference is slight, so "they" are excluded from the supporters.
[d]- One of these is for Dan Aram (7/7), a corpus only present on this one lexeme's plots.

Behaviours of newly introduced non-biblical texts

At the "late end" of the x-axis of the Rezetko-Young plots invariably come Ben Sira and (non-biblical) DSS,[43] both being late in origin. Ben Sira was composed around 180–175 B.C.E. (Kister 2013/I:260), and the DSS manuscripts date from 200 B.C.E. to 68 C.E. (Fassberg 2013/I:663). Being late, one might expect that the proposed LBH features would tend to be close to 100% for these corpora. But, as the "Rel. Freq." columns document in Table 5, they are not. For Ben Sira, the average relative frequency is 6.6%; for DSS, it is 14.9% (being inflated by 10% due to the single 100% finding). These low values appear to rule out fitting all of the text portions with s-curves or, indeed, any monotonic curves. Here are four possible reasons for this behaviour, ranked from least to most likely:

1. Innovation curves are neither monotonic nor monopolistic – the likelihood of this accounting for (m)any of the small relative frequencies is quite low.
2. Too many sample sizes may simply be too small – this may partially account for Ben Sira's behaviour[44] but not for that of DSS.[45]
3. The feature sets for the ten selected lexemes were devised to meet Hurvitz's criterion of appearing in extra-biblical sources – there was no constraint that the late forms monopolise in the time-anchored extra-biblical sources.[46] Hence, the

43 As a plot label, DSS means "non-biblical DSS" (Rezetko and Young 2014:243).
44 The Ben Sira samples have few tokens: 1, 4, 7 (x2), 11, 15, 17, 26 (x2), and 28.
45 The DSS samples tend to be substantial. Counts: 13, 23 (x2), 68, 84, 91, 93, 162, 206, and 212.
46 Remember: Hebrew Ben Sira comes from medieval copies of copies…from the Cairo Geniza.

behaviour may involve non-generalisability.[47]

4. The textual parameters of Ben Sira and DSS may be too different from those of the biblical texts – this may account for the observed behaviour.[48]

Criteria with limited "diachrony community" support

The rightmost column in Table 5 gives my reckoning of the number of "authorities" who actually support a given feature as "late". The lists of candidates are taken from the tables in Young, Rezetko & Ehrensvärd (2008/II:160–214). By way of comparison, 60% of the lexical features in the table have only one supporter, and nearly 80% have only one or two. Almost 25% of the lexical features are supported only by Qimron (Forbes 2012:14–15). My hypothesis is that the fewer supporters a feature has, the less likely it is to be generalisable. The columns in the tables specify eleven authorities, ten if one dismisses Qimron. Three lexical features are supported by nine authorities, three by eight, and three by seven or, leaving out Qimron by eight, seven, or six. Many of the supporters tallied in my Table 5 "support" a given lexeme only to the extent that they supply a reference to previous extensive studies. The level of support – and hence the likely generalisability? – for the ten lexical features selected is low.

Mini-studies: confidence intervals and s-curves

In this sub-section, I examine two Rezetko-Young relative frequency plots in light of CI theory. The plots show how a collection of saturated (100%-value) relative frequencies need not dictate an inference of feature invariance. They also illustrate how easily small changes in method can lead to important-seeming, but illusory, changes in outcome. They, yet again, show how statistical uncertainties due to small sample sizes make shaky intuitive s-curve fitting.

---

[47] On these grounds, one might argue that inclusion of Ben Sira and DSS in the Rezetko-Young plots is an important element of the "*stacked deck*."

[48] As Rezetko and Young put it: "Hebrew inscriptions, the book of Ben Sira, and the non-biblical DSS…are rather inadequate 'anchors' for comparison…because of significant differences related to corpora sizes, subjects, genres, registers, possibly dialects, and so on" (2014:67).

Too small samples

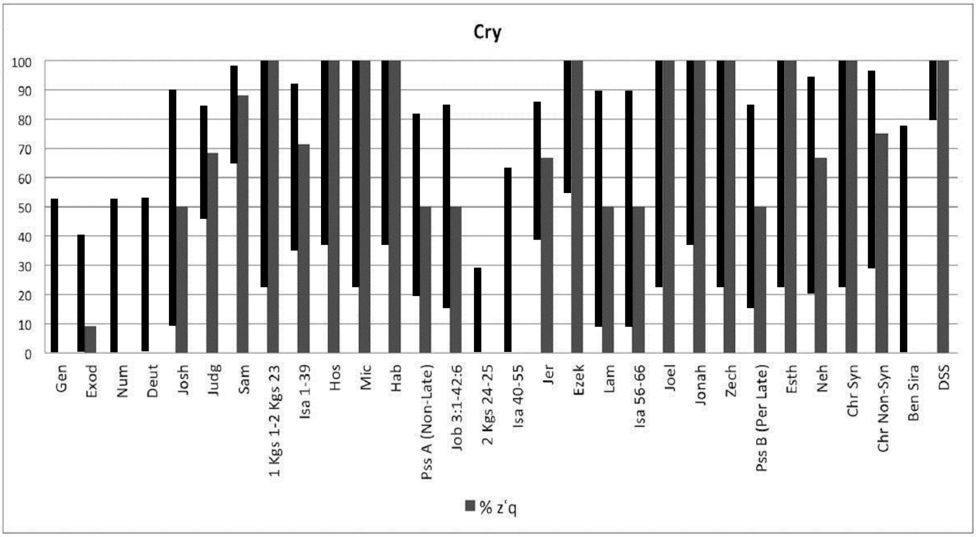

Figure 9. Observed relative frequencies of זעק "cry" across text portions (with 95%-CIs to the left of each relative-frequency bar position)

Figure 9 is Reztko's and Young's Figure 8.10 (2014:281)[49] augmented by the narrow dark CI bars. Among the relative frequency plots for the ten lexemes presented by Rezetko and Young, this plot has by far the largest number of 100%-relative-frequency observations (11 in all). Its sole non-zero observation below 50% is that of Exodus, and there are six 0% observations. All of its thirty sample sizes are under twenty, and more than two-thirds (22) are under five.

Taking the text-portion order adopted by the authors to provide a "conversational framework" (Rezetko and Young 2014:250, n. 17), the observed fluctuations are substantial. Any hope of creating a convincing s-curve based on the raw observations for a single feature like this is highly unlikely. A basically monotonic curve wending its way through the CIs is possible, allowing for the inevitable fluctuations, but these data are simply too few to be statistically trustworthy on their own.

49 Reproduced with permission; the book is freely available at http://sbl-site.org/publications/Books_ANEmonographs.aspx.

Rezetko and Young (2014:399, n. 175) comment that “it hardly needs to be pointed out that none of the line charts in Chapters 8 or 9 resembles an ‘s-curve’ despite organizing the writings in the conventional triparte [sic] periodisation of biblical writings which many Hebraists and biblicists would more or less agree with”. True, few readers would see s-curves lurking in the Rezetko-Young relative frequency plots. However, matters are more complex than the above comment indicates. When they introduce the text portions, Rezetko and Young explain that “[t]he periodisation is organised first according to three historical eras and then according to canon (the latter *not* in a suggested chronological order)” (Rezetko and Young 2014:250; italics in original). Ordering the intra-era texts canonically would be expected to violate historical ordering, as the diachronists would view matters. Hence, it would not surprise them that s-curves are not evident in the plots. There is a further potential obliterator of s-curve behaviour best appreciated by examining Figure 8. There, Holmstedt’s dataset lacks portions with relative frequencies between 53% and 93%. Had he spaced the portions equally along the x-axis, the fitted curve would have exhibited a sudden 40% step between immediately adjacent portions Ps-B and Song, destroying the neat adherence to the posited s-curve.

This suggests a question: What about the behaviour in terms of the three-way division of texts? We begin by examining the behaviour of the era-specific texts for the lexical feature involving זעק, considered in Figure 9 and environs. Using the specifications provided in Rezetko-Young Table 8.1 (2014:249), it is straightforward to use Rezetko-Young Figures 8.9 and 8.10 to work out the relative frequencies for the three sets of texts. The results are tabulated here in Table 6.

Table 6. Incidences of זעק across hypothetical subcorpora

	Incidences (by percent and by count)		
Feature	**Preexilic books**	**Exilic books**	**Postexilic books**
זעק	57% (46/81)	45% (13/29)	76% (25/33)

Table 7. Incidence of זעק across adjusted subcorpora

	Incidences (by percent and by count)		
Feature	**Preexilic books**	**Exilic books**	**Postexilic books**
זעק	51% (46/90)	65% (13/20)	76% (25/33)

"Aha!" declares a revisionist. "The evolution of the feature is not monotonic!" "Wait," says a diachronist, "I detect strange goings on! Consider the characteristics of text portion 2 Kings 24–25. It is a tiny part of the exilic corpus, is absent for six of the eleven relative-frequency plots in Chapter 8, and contains no זעק 'late' tokens."

If we resorb 2 Kings 24–25 into the 1–2 Kings corpus, making Kings whole,[50] then the results for the three sets of text are as shown in Table 7.

"Aha!" says the diachronist. "The feature development is monotonic!"

So, can we decide who is correct? The CIs shown in Figure 10 provide important information. The three intervals on the left pertain to the three corpora proposed by Rezetko and Young; the three on the right, to the restored Kings corpora, after 2 Kings 24–25 has been moved from the exilic to the preexilic corpus. For both the left and right trios, all possible CI pairs overlap. Hence, further testing is required to determine whether the observed relative frequencies differ significantly or not. Applying the MOVER-D algorithm,[51] one finds that the exilic and postexilic values differ significantly for the left trio (HLBH),[52] while the preexilic and postexilic values differ significantly for the right trio (adj HLBH). The other four differences are not statistically significant. With neither grouping are we in a position to claim statistical significance at the 95% level for most of the observed differences.

50 This action belongs to the class "shameless *ex post facto* fiddle".

51 Technical note: The MOVER-D algorithm is discussed in Newcombe (2012:132–138). The algorithm is available at www.medicine.cf.ac.uk/media/filer_public/2013/10/30/mover-d_301013.xls.

52 HLBH = *Historical linguistics and Biblical Hebrew* (Rezetko and Young 2014).

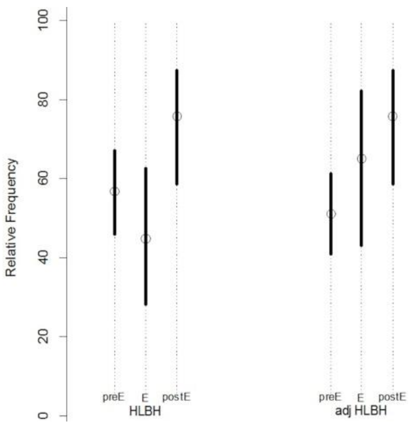

Figure 10. CIs for preexilic, exilic, and postexilic corpora. “adj HLBH” differs from “HLBH” in that 2 Kings 24–25 has been moved from the HLBH hypothetical exilic corpus into the preexilic.

***Piels* of הלך**

Rezetko and Young order the portions corresponding to the dozen most-observed relative frequencies for the *Piel*s of הלך (as per their Figure 8.8) to produce the monotonic curve in their Figure 8.25. The sizes of the right-most dozen texts in their figure are given here in Table 8.

Table 8. Text sizes

1 Kgs1- 2 Kgs 23	Ezek	DSS	Job 3:1- 42:6	Qoh	Isa 55-66	Prov 1-9 & 30-31	Lam	Pss A	Ben Sira	Hab	Pss B
213	59	206	19	28	7	20	5	40	17	3	8

Note that eight of the twelve samples are small (thirty or fewer), with four tiny (fewer than ten).

Figure 11 shows the text portions along with overlaid 95%-CIs. As with the Holmstedt data, the CIs show that the portions for the *Piel*s of הלך might be ordered in many different ways.

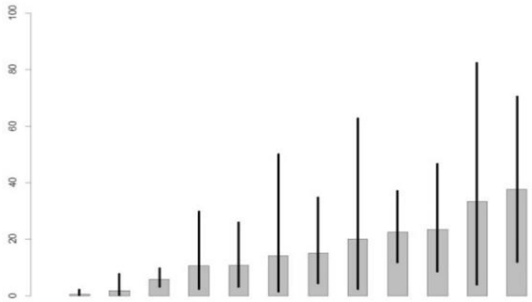

Figure 11. Naïve seriation ordering: *Piel* הלך

CRYSTALLISATION-DATE INFERENCE AND SERIATION

Rezetko and Young (2014:2) are convinced: “The linguistic dating of biblical writings [is] unfeasible”. Odd, methinks. Wasn’t that done for Hebrew spelling back in the early 90s?[53] Granted, orthography is not mainstream – or any-stream – linguistics. But close … and the orthography work suggests that linguistics-based relative dating (“ordered time”) might be doable.

Inverse problems and innovation theory

Indeed, we shall infer interesting, verifiable dating results below. But first, some introductory exposition is needed.

While the relative dating of BH texts may be feasible, its implementation involves overcoming critical impediments. To explain, two important topics need exposition: 1) the possible estimating of crystallisation dates by formulating and solving inverse problems; and 2) the limits and promise of innovation theory for the recovery of crystallisation dates or the like. As we will see, the concept of regularisation that is central for the first topic is critical for understanding and exploiting the second.

Forward and inverse problems

We make a very important distinction at this point, namely that between forward

53 Freedman et al. (1992:125–134). Updated by Forbes & Andersen (2012:127–145).

problems and inverse problems. In non-technical language, forward problems involve moving from known cause(s) to resulting effect(s), while inverse problems involve working backwards from measured effect(s) to inferred cause(s). “Formally, to solve an inverse problem means to discover the cause of a known result” (Tikhonov & Goncharsky 1987:9).

Consider an example from cardiac electrophysiology. It is possible, although highly invasive, to measure in situ the temporal evolution of the complex electrical activity within the heart that leads to its properly sequenced mechanical pumping actions. Given knowledge of the intra-cardiac activation sequence, it is then possible using well-established electromagnetic principles to solve the associated forward problem and compute the evolution of EKG potentials on the body surface. The solution to this forward problem is unique and stable (Tikhonov & Goncharsky 1987:10). However, using measured body surface potentials to infer the underlying cardiac activation sequence, solving the associated inverse problem, is considerably more difficult (Forbes 1974:199–208). This is because, absent special constraints, solutions to inverse problems typically are not stable. When a solution is unstable, a tiny change in a measured value can lead to a huge change in the inverse solution.

Regularisation and innovation theory

Solutions can be stabilised by using special techniques that “sacrifice fit to the data in exchange for solution stability” (Aster et al. 2013:68), a process called regularisation. Regularisation methods “express our prior belief that the type of functions we seek [when solving inverse problems] exhibit a certain type of smooth behaviour” (Hastie et al. 2009:34). At issue is how limiting the regularising constraints should be. Opting for a hard constraint, one might completely specify the function that any solution must approximate. In innovation theory, this function is usually the three-parameter logistic function, a specific form of sigmoid function.[54] Adopting this very strong form of regularisation is probably too restrictive.[55] One mild regularisation requirement is that

[54] For beginners’ information on sigmoid and logistic functions, consult the Wikipedia.
[55] Technical note: Executing logit transformations on the Ellegård “do” data does not linearise it.

solutions comply with some smoothness constraint. We impose this constraint on the seriation solution to which we turn after the following excursus.

An instructive excursus on colonial mortuary art … Really!

Deetz (1996:95) describes the evolution of headstone art in New England over an interval of 140 years:

> Three basic designs were used by the stone carvers of New England between about 1680 and 1820. The earliest of the three is a winged death's-head with blank eyes and a grinning visage. … Sometime during the eighteenth century – the time varies according to location – the grim death's-heads were replaced by a new design, the winged cherub. Toward the end of the eighteenth century the cherubs were replaced in turn by the third basic design, a willow tree overhanging a pedestaled urn.

The three styles prevailed at Stoneham Cemetery during the century-and-a-half under study (Deetz 1996:96). Plotting abundances gives Figure 12 (data from Deetz 1996:97).[56] Note the death's-head style – initially (nearly) monopolising – dwindled and disappeared over the decades. The cherub style appeared, attained a non-monopolising maximum, and then petered out. Meanwhile, the willow-tree-and-urn built up and then achieved near monopoly early in the nineteenth century. The figure visualises the forward problem solution from *decade* to *headstone style*. Question: Suppose you were given the columns of Figure 12 at random, with the date ranges hidden. Could you arrange the columns and recover the correct temporal order? Clearly, you could not, since the four "earliest columns" are identical, as are the two "latest columns." This observation, as well as consideration of possible noise effects, suggests that object classes be ordered based on *multiple* descriptive features, an axiom of pattern recognition theory.

[56] Thanks to Christopher Lowman for alerting me to this dataset. Headstone drawings by Ellen Forbes.

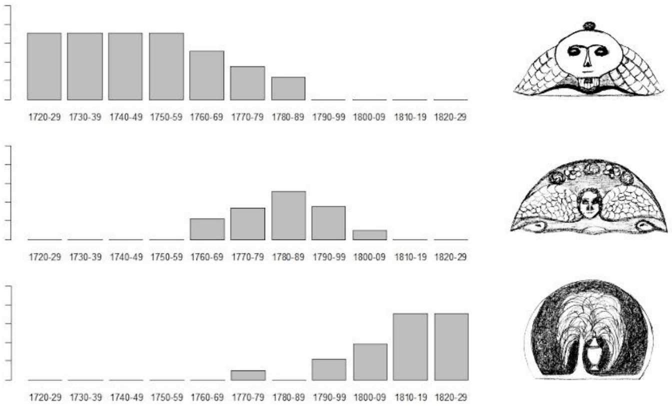

Figure 12. Stoneham headstone designs

Simple seriation: row-shuffling the pictorial abundance matrix

The oldest and simplest embodiment of seriation is called the graphical seriation technique. The basic seriation criterion is as follows:

> Given a matrix of *n* rows and *m* columns of which the elements are nonnegative real numbers, find the permutation of rows that best satisfies this criterion: In each column the elements increase to a maximum and then decrease…in the 'weak' sense [that] two or more consecutive identical elements do not violate the criterion (Marquardt 1978:292).

J. R. Ford showed how to seriate manually (O'Brien & Lyman 1998:121–30). For present illustrative purposes, I use Ford's method. Below, matrix rows correspond to time intervals, columns to linguistic features, and the element at the intersection of a row and a column gives the percentage saturation of the particular feature during the particular time interval.

Having no "crystallisation-dated" BH texts, I use the Ellegård "do" data (Kroch 1989:226, Table 3) for a trial study. Figure 13 shows the data feature-by-feature in "row=percent/column=time" format. Rotating each sub-plot counterclockwise by ninety degrees, aligning the bars with respect to their midpoints, spacing them equally vertically, and combining them appropriately yields the diagram of original data in "row=time/column=feature" format shown in Figure 14.[57]

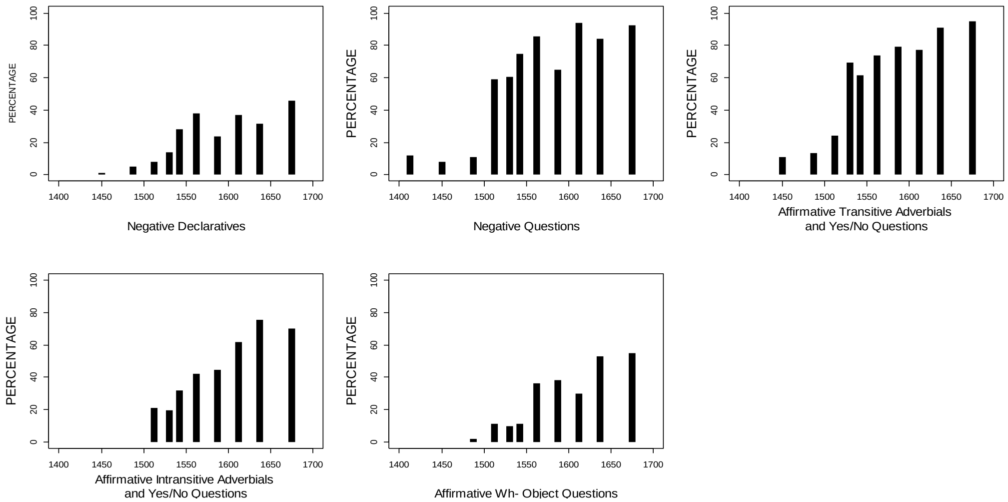

Figure 13. Ellegård's "Do" data (after Kroch)

One proceeds as follows:

1. Identify a tagged corpus for which we know the dates of its constituent portions (row indices) and the extent of saturation (element values) for its descriptive linguistic features (column indices).
2. In the R statistics program:
   a. Create and populate the numerical abundance matrix.
   b. Produce a graphical display wherein the value of each element is shown as a rectangle having a proportionate width.
   c. Before observing the pictorial abundance matrix, randomly permute its rows, keeping the permutation specifics hidden.
   d. Print out the pictorial abundance matrix and cut it into single rows.
3. Manually permute the rows until the row-permuted matrix seems best to meet the seriation criterion specified above.
4. Unseal the permutation details to find the true ordering of the epochs.

57 Figure 14, Figure 15, and Figure 16 each show a pictorial abundance matrix.

The randomly row-permuted data are shown in Figure 15. My seriation-result matrix is Figure 16. Recall that the correctly ordered data are shown in Figure 14.

My ordering differs from the true order only in that I mis-positioned the 1575–1600 row. Based on the Ellegård ‘do’ data, and in spite of its ample noise, *at least the linguistic dating of these English writings by seriation appears feasible.*

The declines in both the “neg. dec.” and “neg. quest.” features between 1550–1575 and 1575–1600 led to my seriation error. Ellegård remarked that the irregularity of the negative question development “is probably due above all to the relatively small number of instances, which exposes the sample to chance variations….[The irregularities in negative declarative corpus development] should however be regarded as significant” (1953:162–163). Kroch theorised: “It seems plausible to hypothesise that the point of inflection in [1550–1575] corresponds to a major reanalysis of the English auxiliary system” (Kroch 1989:227). One declining usage is artifact and the other perhaps genuine.

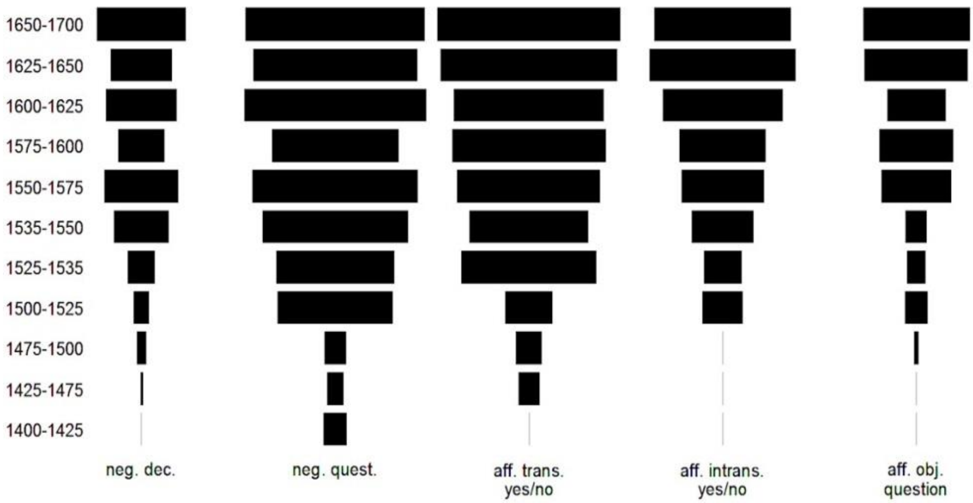

Figure 14. Correct pictorial abundance matrix for Ellegård “Do” data

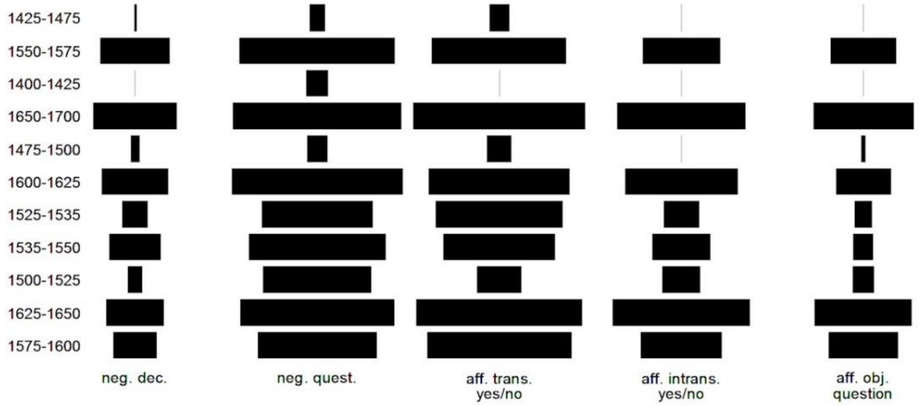

Figure 15. Randomly row-permuted pictorial abundance matrix for Ellegård “Do” data. (When this figure is printed out, the date ranges are hidden.)

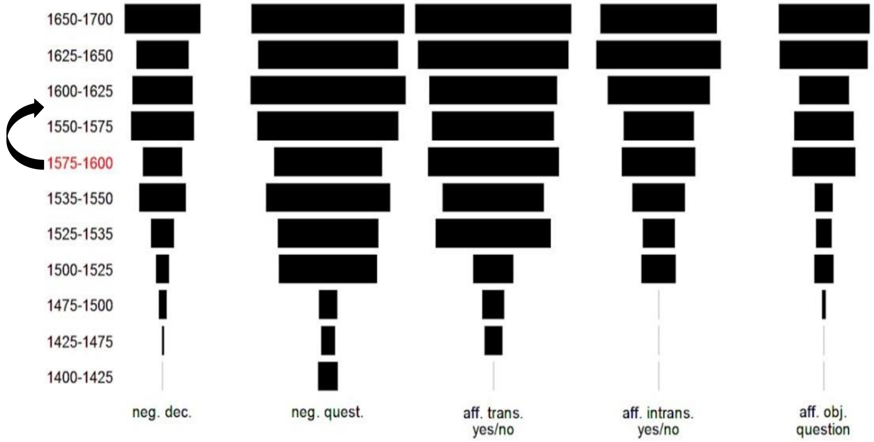

Figure 16. Results for manually-seriated pictorial abundance matrix for Ellegård “Do” data. (The arrow at the left shows the correction needed to make the figure match Figure 14.)

Problems with simple seriation

The manual row-shuffle approach to seriation has several weaknesses: 1) it is not falsifiable in that it always produces a result – be it compelling or feeble; 2) it does not tolerate noise well; 3) it becomes intractable as the number of rows and/or columns gets very large; 4) it does not allow side-information to constrain the solution; 5) it offers no obvious way to allow for additional textual parameters.[58]

Modern seriation: resorting to multivariate methods

There are ways of overcoming these weaknesses, techniques for performing seriation that offer falsifiability, noise tolerance, automated analysis, constraint incorporation, and extensibility. These methods all rely on multivariate statistical procedures.[59]

They typically begin by converting the nxm abundance matrix into an nxn similarity matrix. The elements of the similarity matrix disclose the similarities between all pairs of types, text portions in our case. I used one such multivariate technique ("multidimensional scaling") in my seriation studies for orthography. The elaborate steps of that analysis may be grasped by motivated readers of Freedman, Forbes & Andersen (1992:125–134).

Because of the needs of evolutionary biology, much significant work on seriation has been published in recent years. On the general state of affairs, Gray et al. (2011:1091) remark: "It is ironic that over the past half-century, computational methods in historical linguistics have fallen out of favor while in evolutionary biology computational methods have blossomed". As a result, historical linguistics is the poorer.

We wish to "make the most of data sets composed of a number of related samples of uneven size, …where some of the samples happen to be small" (Robertson 1999:150). Our investigations must use advanced multivariate methods, simple methods being too inflexible.

58 Regarding this final deficiency, Smith and Neiman (2007:47) observed: "If types measure time and, say, social status, then a frequency-seriation diagram will appear messy, or noisy, and deriving a chronological order will be less straightforward". Forbes and Andersen (2012:144) commented on this behaviour.

59 A good grasp of statistics is required to understand multivariate approaches.

EPILOGUE

In this essay, I have explained the concepts of “crystallisation date”, overfitting, non-generalisability, noise types, forward and inverse problems, regularisation, and confidence intervals.

I have proposed several ways of improving accumulation, sometimes quite in agreement with the work of Rezetko and Young (2014) and sometimes at odds. I have also classified, characterised, and suggested ways of taking account of textual parameters that can affect texts. In this, I am again in agreement with Rezetko and Young (2014) as well as several others before them.

I have counselled against viewing textual “authenticity” as crucially dependent on the closeness of texts to spontaneous speech (contra Rezetko and Young 2014). I have warned against assuming a functional form for s-curves (in agreement with Rezetko and Young 2014) and have shown how non-monopolisation, non-monotonicity, and fluctuation seriously compromise the usefulness of simple s-curves. I have used confidence interval theory to show how the fluctuations associated with small samples cause intuitive s-curves to be misleading. I have shown how a small adjustment in the composition of a corpus can alter the ordering of its so-called “EBH-LBH” feature values across sub-corpora.

Most importantly, I have contended that the relative dating of BH texts is not necessarily out of reach (fundamentally at odds with Rezetko and Young 2014). In a pilot study, I have shown how simple seriation can be used to order somewhat noisy Middle-English sub-corpora in time, based on a set of features. Whether innovative seriation methods can be used to advance diachronic analysis is an open question. The extent of the various kinds of noise afflicting our texts may prevent success (consult Forbes 2016). We shall only know whether this is the case after we carry out the analysis.

So, where does all this leave us? Language-based diachrony studies definitely should continue to be actively pursued, incorporating the insights resulting from the diachrony debate and exploiting a statistical technique known as “boosting” – “a procedure that combines the outputs of many ‘weak’ classifiers to produce a powerful 'committee' … a weak classifier [being] one whose error rate is only slightly better than random guessing" (Hastie et al. 2009:337). I believe that significant advances may be within our reach.

"In [a] multitude of counsellors there is safety."

Proverbs 24:6b (KJV).

BIBLIOGRAPHY

Abu-Mustafa, Y, Magdon-Ismail, M & Lin, H-T 2012. *Learning from data: a short course*. N. p.: AMLbooks.[60]

Altmann, G, van Buttlar, V, Rott, W & Strauss, U 1983. A law of change in language, in Brainerd 1983:104–115.

Andersen, F & Forbes, A D 1986. *Spelling in the Hebrew Bible*. Rome: Pontifical Institute Press.

_______ 2012. *Biblical Hebrew grammar visualized*. Winona Lake, IN: Eisenbrauns.

Aster, R, Borchers, B & Thurber, C 2013. *Parameter estimation and inverse problems.* 2nd ed. Waltham, MA: Academic Press.

Bayley, R & Preston, D (eds) 1996. *Second language acquisition and linguistic variation*. Philadelphia: John Benjamins.

Ben Zvi, E, Edelman, D & Polak, F (eds) 2009. *A palimpsest: rhetoric, ideology, stylistics, and language relating to Persian Israel*. Piscataway, NJ: Gorgias.

Berdan, R 1996. Disentangling language acquisition from language variation, in Bayley & Preston 1996:224–230.

Biber, D & Conrad, S 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.

Borges, J 1962. The Babylon lottery, in Borges 1962:65–72.

_______ 1962. *Ficciones*. New York: Grove Press.

Brainerd, B (ed.) 1983. *Historical linguistics*. Bochum: Studienverlag Dr. N. Brockmeyer.

Breiman, L, Friedman, J, Olshen, R & Stone, C 1984. *Classification and Regression Trees*. Belmont, CA: Wadsworth.

Bresnan, J & Hay, J 2008. Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English, *Lingua* 118:245–259.

Brown, L, Cai, T & DasGupta, A 2001. Interval estimation for a binomial proportion, *Stat. Sci.*, 16:101–133.

Capps, E, Page, T & Rouse, W (eds.) 1926. *Josephus*. London: Heinemann.

Carr, D 2011. *The formation of the Hebrew Bible: a new reconstruction*. Oxford: Oxford University Press.

Chambers, J & Schilling, N 2013. *The handbook of language variation and change.* Chichester: Wiley-Blackwell.

Cohen, J 1990. Things I have learned (so far), *American Psychologist* 45:1304–1312.

Conrad, E & Newing, E (eds) 1987. *Perspectives on language and text: essays and poems in*

[60] This book is highly mathematical.

honor of Francis I Andersen's sixtieth birthday. Winona Lake, IN: Eisenbrauns.
Cook, J 2012. Detecting development in Biblical Hebrew using diachronic typology, in Miller-Naudé & Zevit 2012:83–95.
Crystal, D 2008. *A dictionary of linguistics and phonetics*. 6th ed. Malden, MA: Blackwell.
Dana, J & Dawes, R 2004. The superiority of simple alternatives to regression for social science predictions, *Journal of Educational and Behavioural Statistics* 29:317-331.
Davies, P 2003. Biblical Hebrew and the history of ancient Judaic typology, chronology and common sense, in Young 2003:150–163.
de Caën, V 2001. Hebrew linguistics and biblical criticism: a minimalist programme, *JHS* 3:1–30.
Deetz, J 1996. *In small things forgotten: an archaeology of early American life*. New York: Anchor Books.
Duda, R, Hart, P & Stork, D 2001. *Pattern classification*. 2nd ed. New York: John Wiley.
Ellegård, A 1953. *The auxiliary do: the establishment and regulation of its use in English*. Stockholm: Almqvist & Wiksell.
Fassberg, S 2013. Dead Sea Scrolls: linguistic features, in Khan 2013/I:663.
Feldman, L 1992. Josephus, in Freedman 1992/III:985–986.
Firestein, G et al. 2012. *Kelley's textbook of rheumatology*. 9th ed. Philadelphia: Saunders.
Fleischman, S 2001. Methodologies and ideologies in historical linguistics: on working with older languages, in Herring et al. 2001: 33–58.
Forbes, A D 1974. Forward and inverse solutions using distributed current sources, in Rush & Lepeschkin 1974:199–208.
_______ 1987. Syntactic sequences in the Hebrew Bible, in Conrad & Newing 1987:59–70.
_______ 2012. The diachrony debate: perspective from pattern recognition and meta-analysis, *Hebrew Studies* 53:7–42.
_______ 2014. On dating Biblical Hebrew texts: sources of uncertainty/analytic options. Unpublished paper presented at the 2014 ISLP Symposium in St. Petersburg, Russia.[61]
_______ 2016. Crystallisation of Biblical Hebrew texts and diachronic inversion. Unpublished paper presented at the 2016 ISLP Symposium in Stellenbosch.
Forbes, A D & Andersen, F 2012. Dwelling on spelling, in Miller-Naudé & Zevit 2012:127–145.
Freedman, D (ed.) 1992. *Anchor Bible dictionary*. New York: Doubleday.
Freedman, D, Forbes, A D & Andersen, F 1992. *Studies in Hebrew and Aramaic orthography*. Winona Lake, IN: Eisenbrauns.
Gray, R, Atkinson, Q & Greenhill, S 2011. Language evolution and human history: what a difference a date makes, *Phil. Trans. R. Soc. B* 366:1091.
Hand, D 1997. *Construction and assessment of classification rules.* Chichester: John Wiley.
Hastie, T, Tibshirani, R & Friedman, J 2009. *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed. New York: Springer.
Herring, S, van Reenen, P & Schøsler, L 2001. On textual parameters in older languages, in Herring et al. 2001:1–31.
_______ 2001. *Textual parameters in older languages*. Amsterdam: John Benjamins.
Hock, H 2001. Genre, discourse, and syntax in early Indo-European, with emphasis on Sanskrit, in Herring 2001:163–95.
Holmstedt, R 2012. Historical linguistics and Biblical Hebrew, in Miller-Naudé & Zevit 2012:97–124.

[61] Nearing publication by Gorgias.

Hurvitz, A 1972. Linguistic criteria for dating problematic biblical texts, *Hebrew Abstracts* 14:74–79.
_______ 2000. Can biblical texts be dated linguistically? Chronological perspectives in the historical study of Biblical Hebrew, in Lemaire & Saebø 2000:143–160.
_______ 2006. The recent debate on Late Biblical Hebrew: solid data, experts' opinions, and inconclusive arguments, *Hebrew Studies* 47:191–210.
Kallin, J (ed.) 1997. *Focus on Ireland*. Amsterdam: John Benjamins.
Kawashima, R 2013. Stylistics: Biblical Hebrew, in Khan 2013/III:643–650.
Kerby, D 2003. CART analysis with unit-weighted regression to predict suicidal ideation from Big Five traits, *Personality and Individual Differences* 35:249–261.
Khan, G (ed.) 2013. *Encyclopedia of Hebrew language and linguistics*. Leiden: Brill.
Kister, M 2013. Ben Sira, in Khan 2013/I:260–262.
Kroch, A 1989. Reflexes of grammar in patterns of language change, *Language Variation and Change* 1:199–244.
Langbehn, D & Woolson R 1997. Discriminant analysis using the unweighted sum of binary variables: a comparison of model selection methods, *Statistics in Medicine* 16:2679–2700.
Lemaire, A & Saebø, M (eds) 2000. *Congress Volume Oslo 1998.* VTSup 80. Leiden: Brill.
Levin, B 1993. *English verb classes and alternations.* Chicago: University of Chicago Press.
Levin, B & Hovav, M 2005. *Argument realization*. Cambridge: Cambridge University Press.
Mahajan, V & Peterson, R 1985. *Models for innovation diffusion.* Newbury Park, CA: Sage.
Marquardt, W 1978. Advances in archaeological seriation, *Advances in Arch. Methodology and Theory* 1:257–314.
Miller-Naudé, C & Zevit, Z (eds) 2012. *Diachrony in Biblical Hebrew*. Winona Lake, IN: Eisenbrauns.
Montgomery, M 1997. A tale of two Georges: the language of Irish Indian traders in colonial North America, in Kallin 1997:227–252.
Nevalainen, T & Raumolin-Brunberg, H 2003. *Historical sociolinguistics: language change in Tudor and Stuart England*. London: Pearson Education.
Newcombe, R 2012. *Confidence intervals for proportions and related measures of effect*. Boca Raton: CRC Press.
O'Brien, M & Lyman, R 1998. *Seriation, stratigraphy, and index fossils*. New York: Kluwer Academic.
Paolillo, J 2002. *Analyzing linguistic variation: statistical models and methods*. Stanford: CSLI.
Polak, F 2006. Sociolinguistics: a key to the typology and the social background of Biblical Hebrew, *Hebrew Studies* 47:115–162.
_______ 2012. Language variation, discourse typology, the sociocultural background of the Hebrew Bible, in Miller-Naudé and Zevit 2012:301–338.
Polzin, R 1976. *Late Biblical Hebrew: toward an historical typology of Biblical Hebrew prose*. Missoula, MT: Scholars Press.
Qimron, E 1986. *The Hebrew of the Dead Sea Scrolls.* Atlanta: Scholars Press.
Rendsburg, G 2003. A comprehensive guide to Israelian Hebrew: grammar and lexicon, *Orient* 38:5–35.
_______ 2013. Biblical Hebrew: dialects and linguistic variation, in Khan 2013/I:338–341.
Rezetko, R and Young, I 2014. *Historical linguistics and Biblical Hebrew: steps toward an integrated approach.* Atlanta: SBL.
Robertson, I 1999. Spatial and multivariate analysis, random sampling error, and analytical

noise: empirical Bayesian methods at Teotihuacan, Mexico, *American Antiquity* 64/1:150.
Rush, S & Lepeschkin, E (eds) 1974. *Body surface mapping of cardiac fields*. Basel: Karger.
Schneider, E 2013. Investigating historical variation and changes in written documents, in Chambers & Schilling 2013:57–81.
Smith, K & Neiman, F 2007. Frequency seriation, correspondence analysis, and woodland period ceramic assemblage variation in the Deep South, *Southeastern Archaeolog*y 26/1:47–72.
Tikhonov, A & Goncharsky, A 1987. *ill-posed problems in the natural sciences*. Moscow: MIR.
Tov, E 2001. *Textual criticism of the Hebrew Bible*. 2nd ed. Minneapolis: Fortress Press.
_______ 2012. *Textual criticism of the Hebrew Bible*. 3rd ed. Minneapolis: Fortress Press.
Virtanen, T 1992. Issues of text typology: narrative–a 'basic' type of text? *Text* 12:293–310.
Vulanović, R 2007. Fitting periphrastic *do* in affirmative declaratives, *J. Quantitative Linguistics* 14/3:111–126.
Wainer, H 1976. Estimating coefficients in linear models: it don't make no nevermind, *Psychological Bulletin* 83:213–217.
Wang, W & Minett, J 2005. The invasion of language: emergence, change and death, *Trends in Ecology and Evolution* 20/5:263–269.
Watson, W 2013. Poetry: Biblical Hebrew, in Khan 2013/III:151–154.
Young, I 2008. Late Biblical Hebrew and the Qumran Pesher Habakkuk, *JHS*, 8, Art. 25:1–38.
_______ 2009a. Is the prose tale of Job in Late Biblical Hebrew? *VT* 59:606–629.
_______ 2009b. What is 'Late Biblical Hebrew'?, in Ben Zvi et al. 2009:253–268.
Young, I (ed.) 2003. *Biblical Hebrew: studies in chronology and typology*. London: T & T Clark.
Young, I, Rezetko, R & Ehrensvärd, M 2008. *Linguistic dating of biblical texts*. London: Equinox.