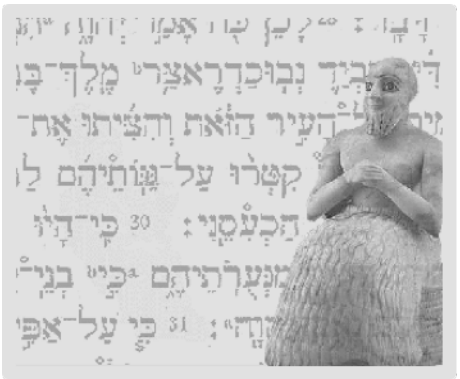
Journal of Hebrew Scriptures

Volume 16, Article 1 DOI:10.5508/jhs.2016.v16.1

An Alternative Approach to the Lexicon of Late Biblical Hebrew

ROBERT REZETKO & MARTIJN NAAIJER



Articles in JHS are being indexed in the ATLA Religion Database, RAMBI, and BiBIL. Their abstracts appear in Religious and Theological Abstracts. The journal is archived by Library and Archives Canada and is accessible for consultation and research at the Electronic Collection site maintained by Library and Archives Canada. ISSN 1203–1542 http://www.jhsonline.org and http://purl.org/jhs

AN ALTERNATIVE APPROACH TO THE LEXICON OF LATE BIBLICAL HEBREW*

ROBERT REZETKO
RADBOUD UNIVERSITEIT NIJMEGEN
& UNIVERSITY OF SYDNEY

MARTIJN NAAIJER
VRIJE UNIVERSITEIT AMSTERDAM

1. INTRODUCTION

The catalyst for the present article and the research behind it was the publication of Avi Hurvitz's *magnum opus*, *A Concise Lexicon of Late Biblical Hebrew*.[1] When we encountered this work several questions came to mind, including What *is* the lexicon of Late Biblical Hebrew (LBH)?, and, What *is late* about the lexicon of LBH? In a separate publication we review the *Lexicon*, discussing its contents and objectives, offering commendations and criticisms, and evaluating closely some (4 of 80) of the individual entries in the book.[2] In this article our objective is to take up the questions we asked above. We want to engage in more depth several significant theoretical and methodological issues related to both the conventional approach and our own tactic toward the study of Biblical Hebrew (BH) language variation and change. We have included a substantial presentation and discussion of recent research undertaken by us, which aims to characterize the lexicon of LBH, and it is intended to function both as a further evaluation of the *Lexicon* and as an illustration of what a new approach can offer to the ongoing debate about the history of BH.

* We thank Dirk Bakker, Martin Ehrensvärd, Marianne Kaajan, Ian Young, and several anonymous reviewers for their helpful remarks on various drafts of this article.

[1] Avi Hurvitz, in collaboration with Leeor Gottlieb, Aaron Hornkohl, and Emmanuel Mastéy, *A Concise Lexicon of Late Biblical Hebrew: Linguistic Innovations in the Writings of the Second Temple Period* (VTSup, 160; Leiden: Brill, 2014).

[2] Robert Rezetko and Martijn Naaijer, Review of Avi Hurvitz *et al.*, *A Concise Lexicon of Late Biblical Hebrew: Linguistic Innovations in the Writings of the Second Temple Period*, *JHS* 16 (2016). The review can be accessed at http://jhsonline.org/reviews_vol.html.

In conventional literature related to diachrony in BH, like the *Lexicon*, scholars collect late variants and their early alternatives as if they all have the same value. Therefore the output of Hurvitz's method of linguistic dating is simply an integer, which represents the number of different late variants in a given biblical book or text segment, so every linguistic variant has the same weight, namely 1. In the following analysis we want to demonstrate that this approach is an oversimplification, by showing the variety of distributional patterns of the early and late variants in a selection of books in the Bible.[3] We do this by studying the distribution of the late variants cited in the *Lexicon*, *together with their early alternatives*. We believe that this approach can lead to a better understanding of the heterogeneity of the material in the *Lexicon* and in the corpus of late biblical writings.

The books in which we study the variables[4] of the *Lexicon* are the (so-called) Transitional BH (TBH) and LBH books of Ezekiel, Haggai, Zechariah, Malachi, Song of Songs, Qoheleth, Esther, Daniel, Ezra, Nehemiah, and Chronicles.[5] In the case of Chronicles we study the attestation of variants in the complete book, but we also look at the distribution of variants in its non-synoptic and synoptic portions. Our aim is to get a clear and complete picture of the distribution of the variants in the selected books by displaying the results visually.

All the data required to examine and recalculate our results, together with all our calculations, can be found at GitHub,

---

[3] Note that even though we talk here about "early and late variants" we do not necessarily agree that the so-called late variants are actually relatively or absolutely late, and vice versa (cf. Rezetko and Naaijer, "Review of Avi Hurvitz *et al*.," §6.7). Nevertheless, throughout this article we have refrained from systematically putting quotation marks around the words "early," "earlier," "old," "late," "latest," "new," "transition," "transitional," "preexilic," "exilic," "postexilic," and so on.

[4] A *variable* is the general or abstract idea, such as "kingdom," whereas the actual instantiations of the variable are the *variants*, for example (early) מַמְלָכָה and (late) מַלְכוּת. For discussion of this and related terminology see Robert Rezetko and Ian Young, *Historical Linguistics and Biblical Hebrew: Steps Toward an Integrated Approach* (SBLANEM, 9; Atlanta: SBL Press, 2014), 46–7.

[5] We decided to study the language of only these eleven late (exilic and postexilic) books for two reasons, first for the sake of manageability, and second because our focus, like the *Lexicon*'s, is mainly on the lexicon of LBH. Therefore our study excludes the books of the Pentateuch and Former Prophets, most of the Latter Prophets, and several of the Writings. However, for the purpose of illustration and to further the discussion we examine the distribution of four variables in the entire MT Bible in our separate review of the *Lexicon* ("Damascus," "after that," "end," Babylonian months) (Rezetko and Naaijer, "Review of Avi Hurvitz *et al*.," §6).

https://github.com/MartijnETCBC/ReviewLexiconOfLBH;[6] in addition, the data are also available at DANS, http://dx.doi.org/10.17026/dans-256-4hcy. We intend to give a balanced picture of linguistic variation and continuity in BH based on the variables in the *Lexicon*, but we also know that the following analyses do not cover all the relevant problems and challenges. Therefore, we invite everyone who wants to continue working with these data to copy the spreadsheet, add more features to it, and process them.

Our main objective in this article is to present an alternative, quantitative approach to describing the lexical stock of late biblical writings. While we limit our investigation to the eighty words and expressions in the *Lexicon*, our method could be used for studying many other grammatical and lexical features of BH, or the language of any other corpus of writings for that matter. Our line of attack is descriptive, with emphasis on transparency and thoroughness. Our main conclusion is that the incidence of late language in late biblical writings is rare and idiosyncratic. The eighty late words and expressions in the *Lexicon* constitute only a very thin layer of the vocabulary of late writings. And as a matter of fact the relative scarceness or complete absence of many late variants in late writings stands in stark contrast to what the traditional view of early and late BH has led to believe. For example, the common idea that late writers often elected to write their compositions using late words instead of early ones is fundamentally flawed. We elaborate on this result further along in this article. One important implication of this outcome (i.e., late language is rare and idiosyncratic in late biblical writings), is that the value of relatively sporadic late language for linguistic periodization and linguistic dating has been overemphasized.

2. THEORETICAL FRAMEWORK

The theoretical framework of our approach to the lexicon of LBH originates in variationist sociolinguistics.[7] In a nutshell its aim is to describe and explain differences (variation) and similarities (continuity) between distinctive spoken or written specimens in regard to

[6] In this repository one can find the datasets and the Python and R scripts we used for our analyses.

[7] For a detailed introduction with extensive citation of literature see Rezetko and Young, *Historical Linguistics*, 211–43. Another significant application of variationist sociolinguistics to BH is Dong-Hyuk Kim, *Early Biblical Hebrew, Late Biblical Hebrew, and Linguistic Variability: A Sociolinguistic Evaluation of the Linguistic Dating of Biblical Texts* (VTSup, 156; Leiden: Brill, 2013). Among the many general publications on variationist sociolinguistics, see Sali A. Tagliamonte, *Analysing Sociolinguistic Variation* (Key Topics in Sociolinguistics; Cambridge: Cambridge University Press, 2006); idem, *Variationist Sociolinguistics: Change, Observation, Interpretation* (Language in Society, 40; Chichester: Wiley-Blackwell, 2012).

"competing" forms or uses with the same meaning (that is, different ways of saying the same thing). For example, to what extent do some (late) biblical writings select late סוֹף ("end") in comparison with others that select early alternatives (אַחֲרִית, קֵץ, קָצֶה, קְצָה) and still others that select some combination of both the old and new words?[8] Our approach to this question is bottom-up, data-driven, and descriptive. We aim to present all the linguistic facts clearly and fully. We ignore conventional presuppositions about dates of biblical writings and periods of BH. All of this is in contrast to the top-down, theory-driven, and prescriptive (linguistic dating) approach which we regard as circular because in the end it simply "proves" what it assumes.

3. DATA COLLECTION AND FORMAT

From each of the eighty entries in the *Lexicon* the late variants and their early alternatives were extracted, as much as possible (cf. below), as they are found in the *Lexicon*. Then, for each of the words and expressions, we recorded all the verses in which the late variants are found and all the verses in which the early alternatives are found in the core LBH books (Esther–Chronicles) and in books sometimes characterized as TBH or LBH (Ezekiel, Haggai–Malachi, Song of Songs, and Qoheleth[9]). The data were collected manually and subsequently stored in a spreadsheet. For each occurrence of the variables in the *Lexicon* the following information was recorded in individual columns:

- *Enumeration*: This integer variable signifies the row number of each token in the spreadsheet. Late variants and early alternatives are enumerated separately.
- *Hur_nr* ("Hurvitz number"): This integer variable indicates the place of the feature in the *Lexicon*.[10]

---

[8] Hurvitz, *Lexicon*, 188–90; cf. Rezetko and Naaijer, "Review of Avi Hurvitz *et al.*," §6.4.

[9] Our choice not to include the book of Qoheleth in the group of core LBH books is based on the *Lexicon*: "The linguistic character of LBH is thus established first and foremost on the basis of features documented exclusively, or predominantly, in compositions from the first group. Data derived from books in the second category must be treated with caution" (Hurvitz, *Lexicon*, 4).

[10] 1. אַגֶּרֶת; 2. אֶדֶר; 3. אֲדַרְכֹן; 4. אַחַר זֶה / זֹאת (כָּל); 5. אֵין + inf; 6. אַחֲרֵי; 7. אֱלוּל; 8. √בהל; 9. בֵּץ, בּוּץ; 10. בִּזָּה; 11. בִּירָה, בִּירָנִית; 12. אֶרֶץ יִשְׂרָאֵל; 13. בֵּית (הָ)אוֹצָר; 14. בֵּית הַכִּפֻּרֶת; 15. בֵּית מִקְדָּשׁ; 16. בֵּית קִבְרוֹת; 17. בֵּית קֹדֶשׁ הַקָּדָשִׁים; 18. √בעת; 19. בַּקָּשָׁה; 20. בָּרוּךְ אַתָּה ה׳; 21. גַּזְבָּר; 22. גַּנָּז; 23. גִּנְזָךְ; 24. √דוח; 25. דָּוִיד; 26. דַּרְכְּמוֹן; 27. דַּרְמֶשֶׂק; 28. √דרש; 29. דַּת; 30. זָוִית; 31. (ה)טוב(ל) + √זכר; 32. √זמן; 33. זְמָן; 34. זַן; 35. √חפה; 36. טֶבֶת; 37. √יחש; 38. יַחַשׁ; 39. יְרוּשָׁלַיִם; 40. יֹאשִׁיָּה; 41. √ישט; 42. כְּאַחַד; 43. כִּסְלֵו; 44. √כעס; 45. כִּרְצוֹן (√עשה); 46. כְּתָב; 47. לְאֵין; 48. (מְאֹד) לְהַרְבֵּה; 49. לָרֹב מְאֹד;

- *Root*: This is a string variable which is a transcription of the Hebrew word(s).[11]
- *Hur_late* ("Hurvitz late"): This binary variable indicates whether it concerns a late variant, with value 1, or one of its early alternatives, with value 0.
- The columns *Book* (string), *Chapter* (integer), and *Verse* (integer) indicate where the variant can be found in the Masoretic Text (MT) of the Bible.
- *Occu_verse* ("Occurrences per verse"): This integer variable shows the number of times the variant occurs in the verse.
- *ChrS_ChrN* ("Chronicles Synoptic—Chronicles Non-Synoptic"): This is a string variable which has the value "ChrS" if the variant is a synoptic example in the book of Chronicles and "ChrN" if the variant is a non-synoptic example in the book of Chronicles.[12]
- *Notes*: This column has optional notes.

We must make several important clarifications in regard to our collection of the data. First, the *Lexicon* does not provide full distributional data for the early alternatives.[13] This means that we have had to search and record those ourselves. Second, the *Lexicon* does not argue for the semantic equivalence of the early and late variants.[14] Sometimes a suggested early variant may not be semantically equivalent to a late one; sometimes some tokens of an early variant

---

50. וּלְמָעְלָה . . . מ; 51. מַאֲמָר; 52. מַדָּע; 53. מִדְרָשׁ; 54. מַלְכוּת; 55. מְעַטִּים; 56. מַעֲרָב; 57. נְבוּאָה; 58. √נדב; 59. נִיסָן; 60. אִשֶּׁה + √נשׂא; 61. סוֹף; 62. סִיוָן; 63. סֵפֶר מֹשֶׁה; 64. . . . עַד לְ; 65. עֲזָרָה; 66. פִּתְגָם; 67. צוּרָה; 68. צָפִיר; 69. צַרְךָ; 70. √קבל; 71. √קום; 72. √קרה; 73. רָאוּי; 74. שְׁבָט; 75. √שלט; 76. שִׁלְטוֹן; 77. שַׁלִּיט; 78. תַּכְרִיךְ; 79. תַּלְמִיד; 80. תַּעֲנִית.

[11] In the database we have adopted a slightly modified version of the transliteration scheme used for the database of the Eep Talstra Centre for Bible and Computer: A (*aleph*), B, G, D, H, W, Z, X (*khet*), V (*tet*), J (*yod*), K, L, M, N, S, E (*ayin*), P, Y (*tsadè*), Q, R, F (*sin*), C (*shin*), T.

[12] To elaborate further, the classification of Chronicles is based on synoptic and non-synoptic items, not passages. In other words, synoptic items are parallel to the same items in the corresponding passages in Samuel–Kings, and non-synoptic items are not parallel to the same items in the corresponding passages in Samuel–Kings. For example, in the synoptic passage 1 Kgs 2:46b–3:15 // 2 Chr 1:1–13, defective דָּוִד occurs in 1 Kgs 3:1, 3, 6, 7, and plene דָּוִיד in 1 Kgs 3:14; 2 Chr 1:1, 4 (x 2), 8, 9. Although the passage is synoptic, only one instance of "David" stands in parallel, in 1 Kgs 3:6 // 2 Chr 1:8. However, there Kings has דָּוִד and Chronicles has דָּוִיד. Therefore in our analysis דָּוִיד in 2 Chr 1:8 is regarded as a non-synoptic or non-shared token of late דָּוִיד (just like the other examples in 2 Chr 1:1, 4 [x 2], 9).

[13] Rezetko and Naaijer, "Review of Avi Hurvitz *et al.*," §5.6.1.

[14] Ibid., §5.6.6.

may not be semantically equivalent to a late one; and sometimes an early variant may itself have undergone semantic change. Our policy has been to accept the *Lexicon*'s judgments about early and late variants, and we have tried to discard illegitimate or uncertain tokens. Third, despite the preceding caveats, we are the first to admit that there is an inevitable aspect of subjectivity in the inclusion/exclusion of particular early words/expressions and tokens. Undoubtedly we will have made some mistakes. Nevertheless, we believe that the substantial amount of data we are examining diminishes the significance of those mistakes. Furthermore, all the relevant data are publicly available and anybody is free to check them themself. Fourth, and finally, in our analysis of the data we have tried to stay as close to the *Lexicon* as possible. However, in the case of item 7 ("the land of Israel") the early alternatives are difficult to count, because there are some subtle semantic nuances. The spreadsheet contains the cases of the early alternatives of item 7 in 1 Chronicles, but we decided to exclude the feature from the analyses. Therefore the analyses of the data are based on seventy-nine instead of eighty variables.

4. METHOD

4.1. ATTESTATION OF EARLY AND LATE VARIANTS IN COMPLETE BOOKS

To begin we want to find out which late variants occur in the books characterized as TBH and LBH and what the proportion of late variants versus early alternatives is. It is important to know this proportion, because it gives insight into the notion of replacement.[15] First, for each of the books under consideration, the absolute amount of all the late variants and their early alternatives are counted and represented in a back-to-back barplot (figures 3–13 in §5.2). Second, for each of the books under consideration, the fraction of late attestations is calculated by dividing the absolute amount of occurrences of each late variant by the absolute amount of occurrences of each late variant and its early alternative(s) combined (figures 18–28 in §5.2). These analyses are done also for the core LBH books combined (figures 16, 31) and for non-synoptic and synoptic portions of Chronicles (figures 14–15, 29–30).[16]

[15] The notion of "replacement" or (absolute) "contrast" between early and late writings has special significance in Hurvitz's method. See Rezetko and Naaijer, "Review of Avi Hurvitz *et al.*," §5.7.5.3 and §5.7.5.4.

[16] See n. 12 on the meanings of synoptic and non-synoptic Chronicles in the present study.

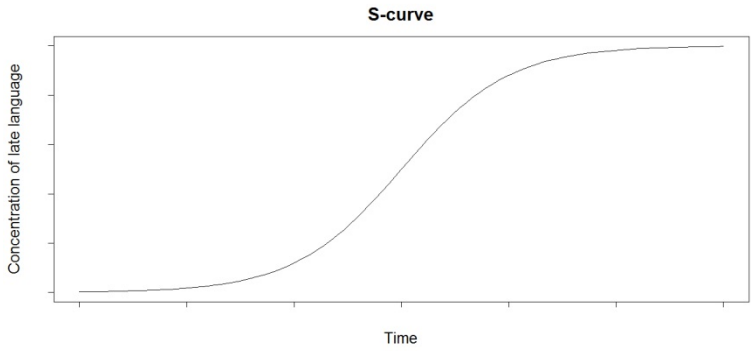
4.2. DIFFUSION OF EARLY AND LATE VARIANTS IN THE CORPUS OF TBH AND LBH BOOKS

An important issue in historical linguistics is the diffusion (or spread) of linguistic innovations. Here we have chosen to study the diffusion of the late variants and their early alternatives throughout the TBH and LBH books. This is done in a very simple way. For all the late variants and their early alternatives reported in the *Lexicon*, the number of books in which they occur are counted and for each of the eleven selected books we display the number of books in which the late variants and their early alternatives are found (figures 32–42 in §5.3). A smoothing line indicates the general tendency in the data. The results are also presented for all the selected books combined (figure 43).

4.3. REPLACEMENT VERSUS VARIATION

In various recent publications the S-curve has been used as a tool for plotting diachronic variation in BH.[17] Though we see the S-curve as an illustrative tool and reject its diagnostic use in these publications,[18] we believe it has some interesting useful properties. The following figure illustrates the S-curve:

Figure 1



The curve shows a *hypothetical* (or *idealized*) gradual replacement of a certain linguistic variant in the course of a certain time span. At the start of the curve the late variant under consideration is absent and only its early alternative is present in a certain text or corpus. At the end of the curve the early alternative is replaced completely and

[17] For example, in John A. Cook, "Detecting Development in Biblical Hebrew Using Diachronic Typology," in Cynthia Miller-Naudé and Ziony Zevit (eds.), *Diachrony in Biblical Hebrew* (LSAWS, 8; Winona Lake, IN: Eisenbrauns, 2012), 83–95; Robert D. Holmstedt, "Historical Linguistics and Biblical Hebrew," in Cynthia Miller-Naudé and Ziony Zevit (eds.), *Diachrony in Biblical Hebrew* (LSAWS, 8; Winona Lake, IN: Eisenbrauns, 2012), 97–124.

[18] Rezetko and Young, *Historical Linguistics*, 223–6, 233–40.

only the late variant occurs. In the time between the initial situation and the complete replacement there is a situation in which both variants can be found in varying proportions. In terms of variation the following pattern can be observed: at the start of the curve there is no variation, then the variation starts to increase as the late variant begins to replace the early alternative, halfway through the variation is at its maximum, and after this maximum the variation begins to decrease until it is low again after the (perhaps complete) replacement.[19]

In the following calculation we want to get an impression of both the rate of replacement and the rate of variation within all the biblical books under consideration. In the framework of the conventional diachronic model of BH we expect that early biblical books will have a low rate of replacement and a low rate of variation, late biblical books will have a low rate of variation but a high rate of replacement, and transitional biblical books will have a high rate of variation and a rate of replacement that is more or less halfway.

For each of the eleven TBH and LBH books the rates of replacement and variation are measured. The rate of replacement is measured in each book for each variable that occurs in it. The rate of replacement of a specific variant in a specific book is simply the late fraction, as explained in §5.2. The variation is measured by calculating the entropy of each variable in each book. The entropy is a measure of disorder or chaos in a system, and here we use it to determine how much disorder there is in the use of early and late variants in BH.[20] The entropy of a system can be calculated with the formula

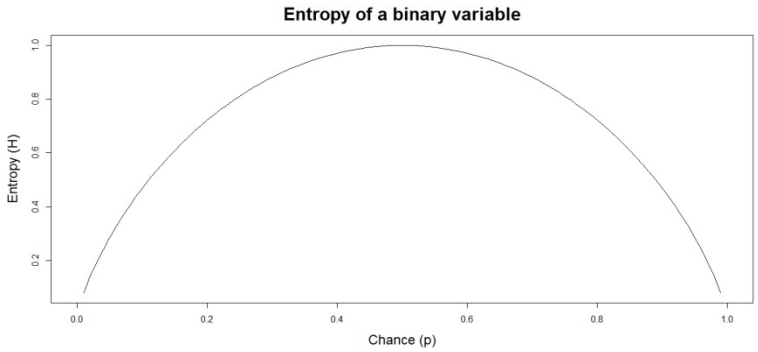
$$H(p) = -p\log_2(p) - (1-p)\log_2(1-p)$$

in the case of a binary variable; p is the chance of the first variant, 1-p is the chance of the second. The entropy H has a value between 0 and 1. The following figure shows the curve of H as a function of p in the case of a binary variable:

[19] For a more nuanced discussion of the S-curve we refer again to ibid., 223–6, 233–40.

[20] For an explanation of the use of entropy in linguistics see Christopher D. Manning and Hinrich Schütze, *Foundations of Statistical Natural Language Processing* (Cambridge, MA: MIT Press, 1999), 60–2.

Figure 2



In this formula p is the chance of finding in a book a late variant of a certain variable. For instance, item 25 (the spelling of the name "David") can be found only in the late form in the book of Chronicles. In that case, p has the value 1, which results in an entropy of 0. In the same book, item 42 ("together") can be found once in the late form and once in the early form, so p is 0.5 in this case and the entropy of this variable is 1. For each of the variables in each book the entropy is calculated. Then the average entropy of each biblical book is calculated by summing the entropies of separate variables and dividing the result by the total number of variables from the *Lexicon* in the biblical book at hand.

5. RESULTS

5.1. INTRODUCTION

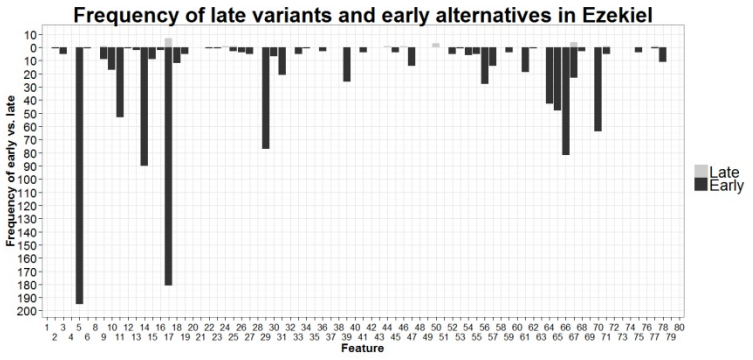
In the following sections the results of the experiments described above are presented. We begin by presenting, first, the frequency of late variants versus early alternatives and, second, the fraction of late tokens. For the sake of illustration, and because it is unnecessary and impractical to discuss each of the eleven books we have studied, we have chosen to discuss the results for the books of Qoheleth and Ezra separately, after which we offer a general impression of the complete results.

5.2. ATTESTATION OF EARLY AND LATE VARIANTS IN COMPLETE BOOKS

The results of the analysis of the late variants and their early alternatives in complete books and non-synoptic and synoptic Chronicles (cf. §4.1) can be found in the following figures. For each book in figures 3–15 all the variables attested are displayed (ordered by feature number), with the absolute number of late attestations in the upper half and the absolute number of their early alternatives in the lower half. When there are two or more early alternatives of one late variant, their values are added together. We believe that this policy is justifiable in the present context, since it is similar to

the approach taken in the *Lexicon*, but in general and in a different kind of study each variant should be individually weighed.[21] For each book in figures 18–31 the fraction of late variants is displayed (ordered firstly by frequency and secondly by feature number). Only those variables are displayed that can be found in the early and/or late variant(s) in the book under consideration. If the value of the fraction of a certain variable is 1, this means that only the late variant can be found in this book, and if the value is 0, then only the early alternative is present. Figures 16 and 31 show the results for the core LBH books combined. Figure 17 shows the concentration of early and late tokens in each of the TBH and LBH books.

Figure 3



[21] In some cases it could also be better to group features together, such as the Babylonian month names, because they form a natural system.

Figure 4

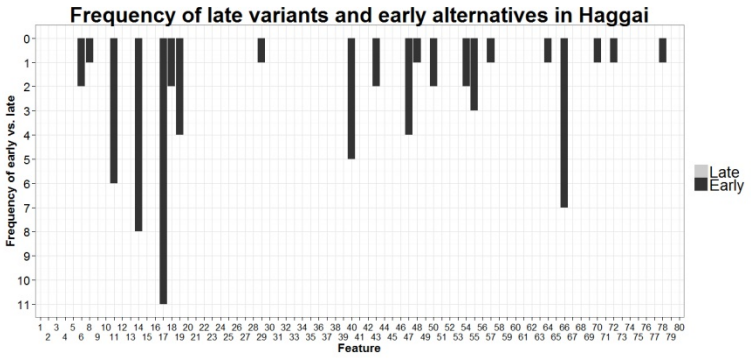


Figure 5

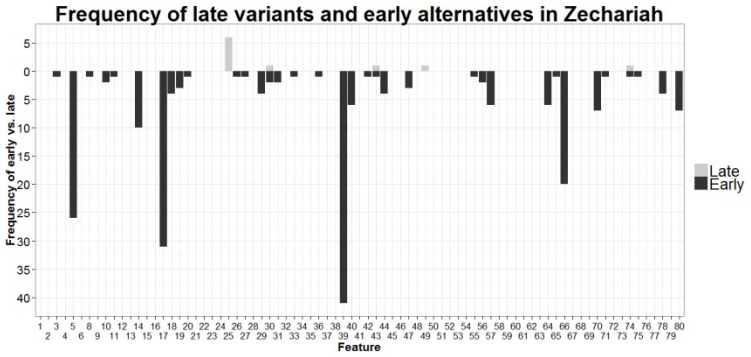


Figure 6

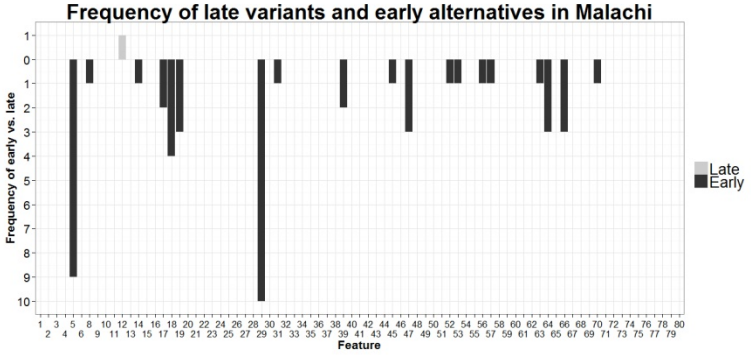


Figure 7

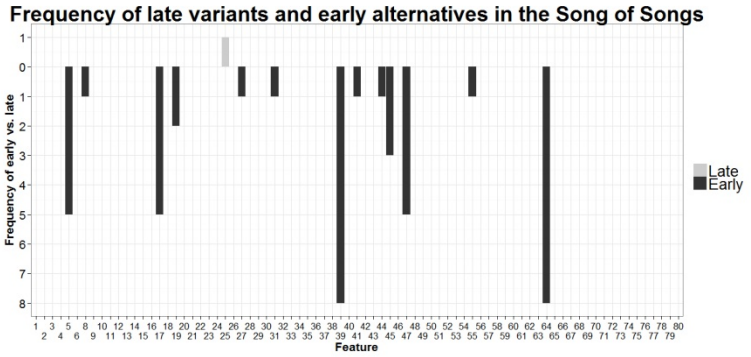


Figure 8

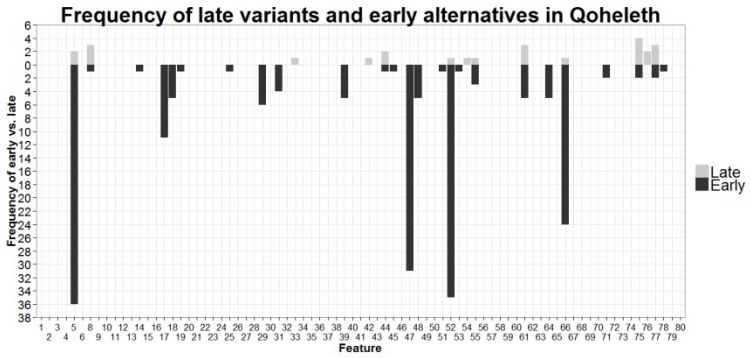


Figure 9

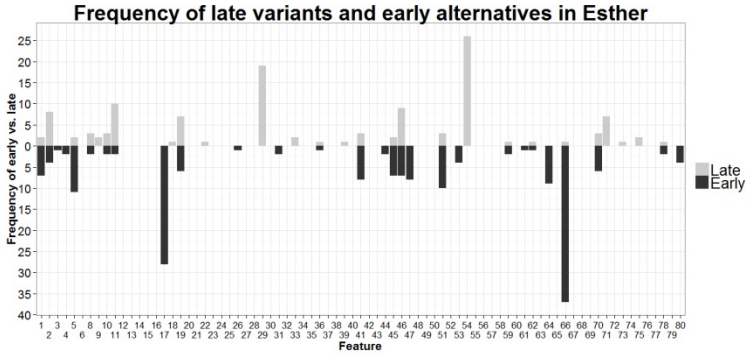


Figure 10

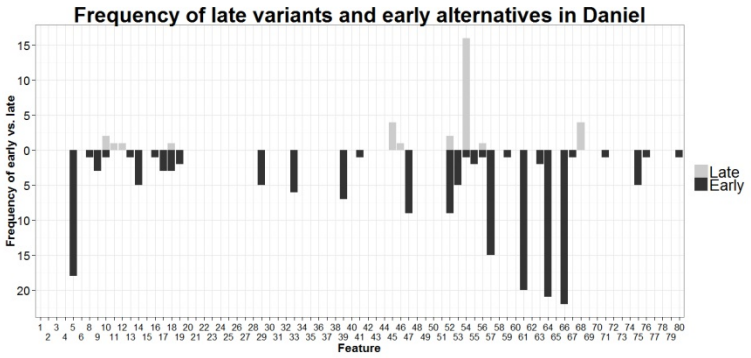


Figure 11

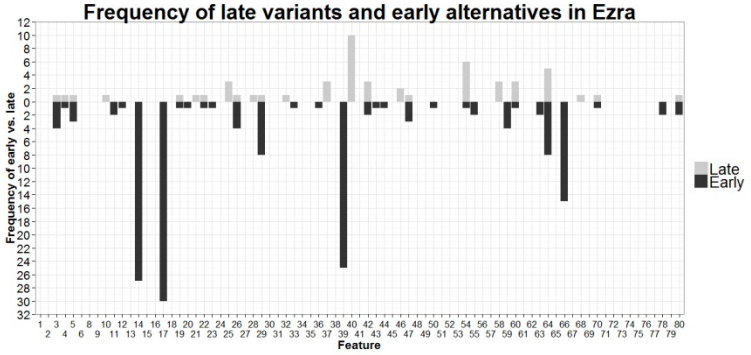


Figure 12

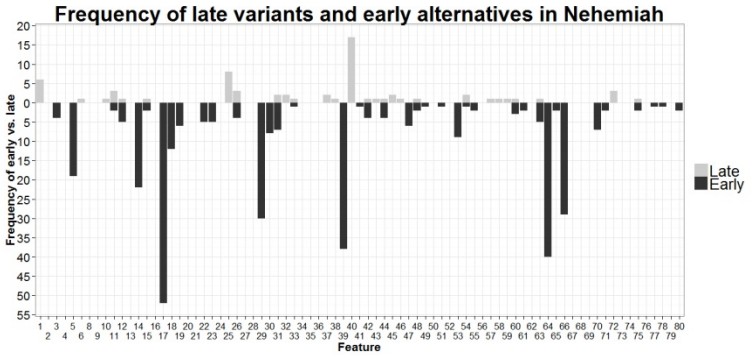


Figure 13

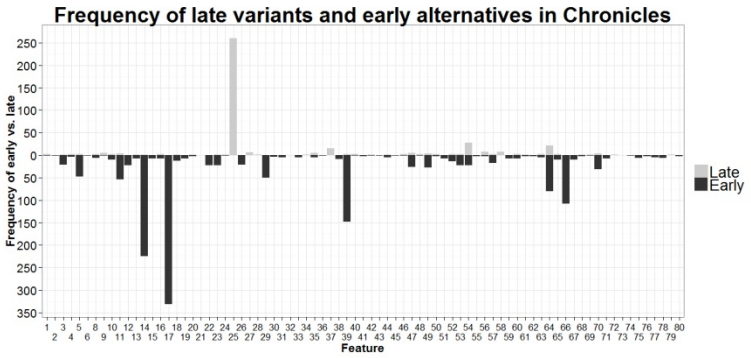


Figure 14

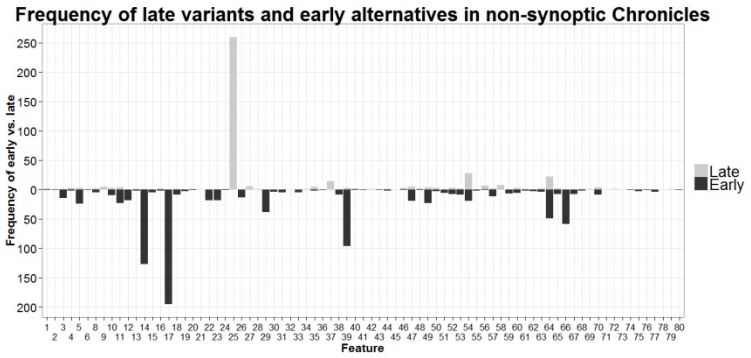


Figure 15

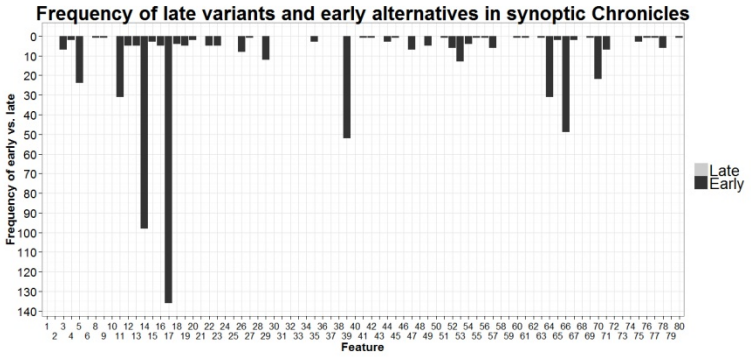


Figure 16

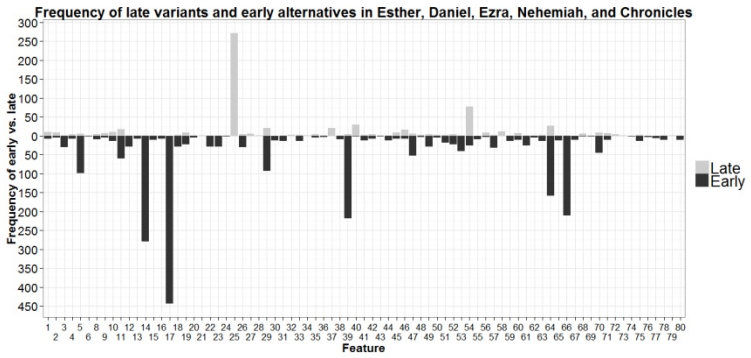


Figure 17

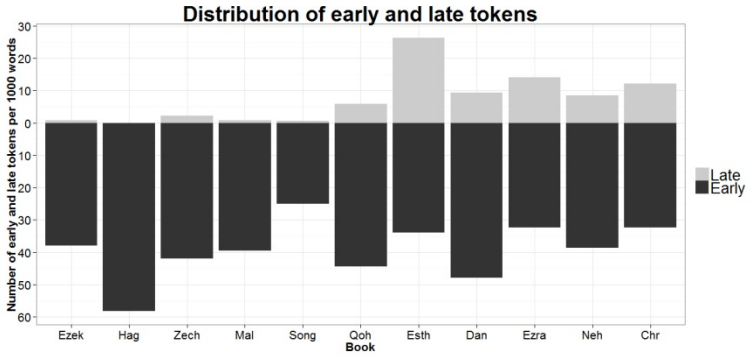


Figure 18

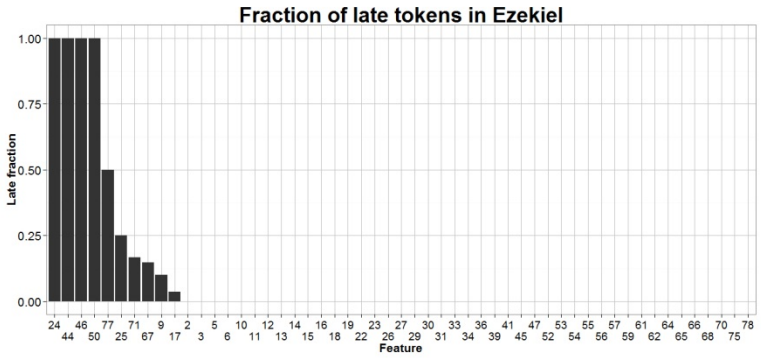


Figure 19

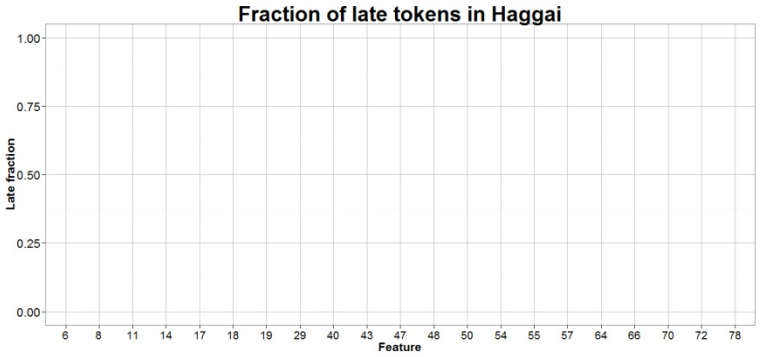


Figure 20

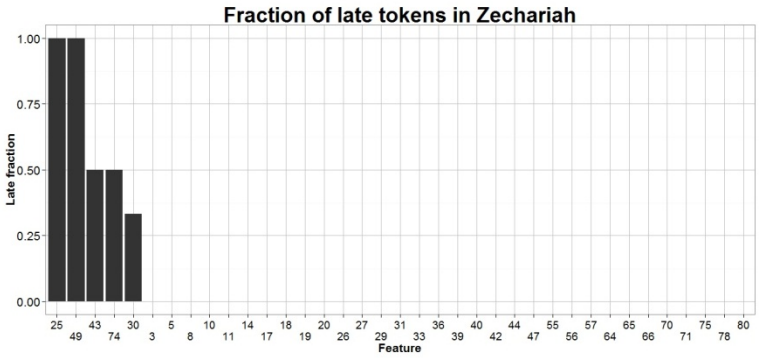


Figure 21

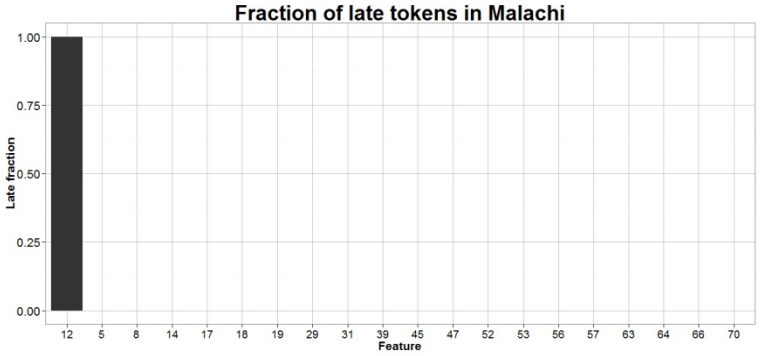


Figure 22

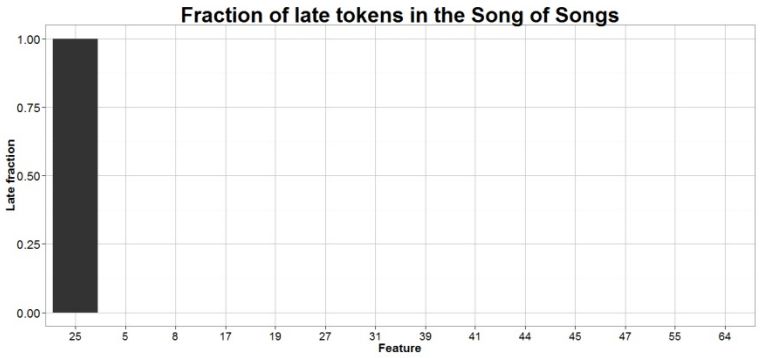


Figure 23

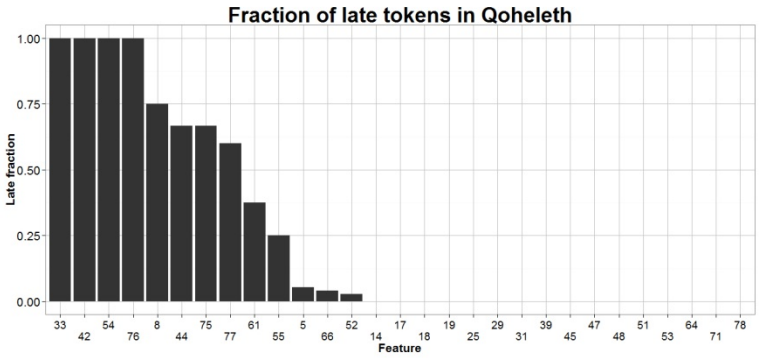


Figure 24

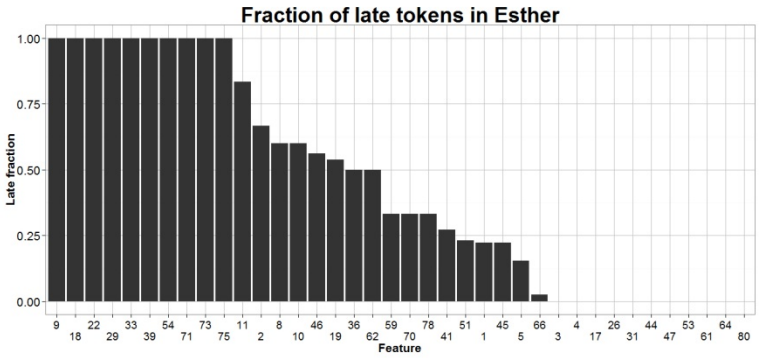


Figure 25

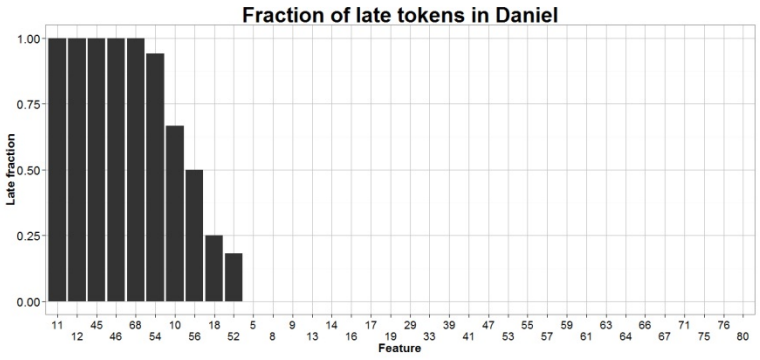


Figure 26

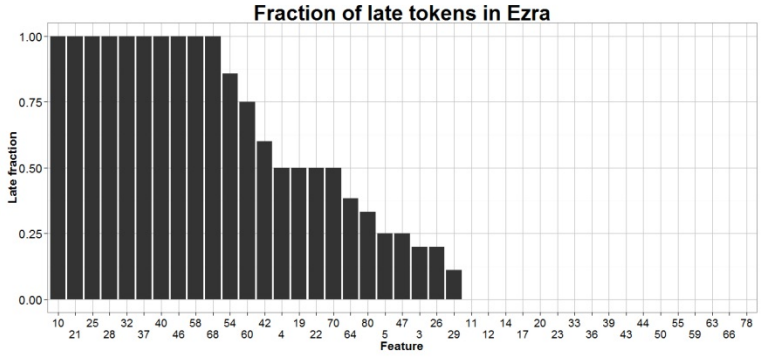


Figure 27

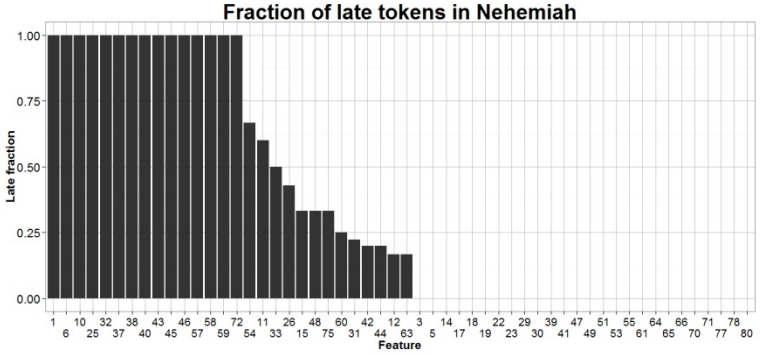


Figure 28

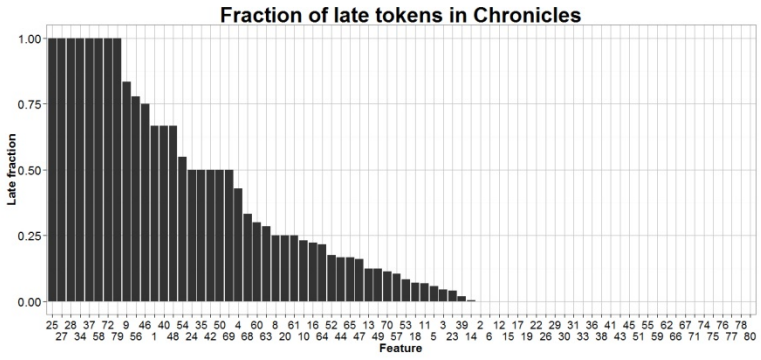


Figure 29

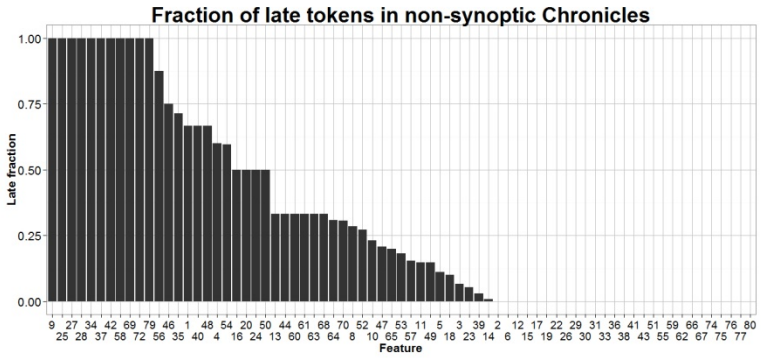


Figure 30

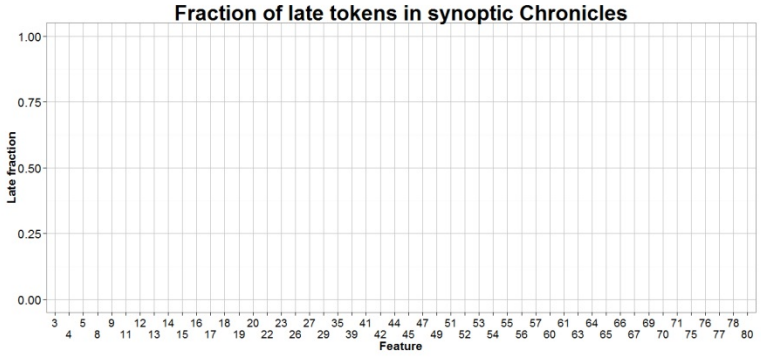
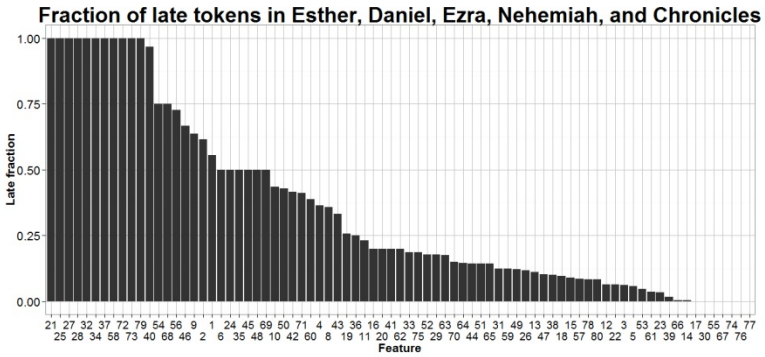


Figure 31



As mentioned above, first we discuss the books of Qoheleth and Ezra and then we give a general impression of the complete results.

Concerning the book of Qoheleth, figures 8 and 23 show that there is a relatively low amount of late language in the book. There are only four variables for which the book contains the late variant exclusively (items 33, 42, 54, 76), but there are sixteen variables for which it has the early variant exclusively (items 14, 17, 18, 19, 25, 29, 31, 39, 45, 47, 48, 51, 53, 64, 71, 78). For nine variables it contains both the early and late variants (items 5, 8, 44, 52, 55, 61, 66, 75, 77).

The book of Ezra shows a slightly different pattern. Figures 11 and 26 show that for ten variables the book contains the late variant exclusively (items 10, 21, 25, 28, 32, 37, 40, 46, 58, 68), but for seventeen variables it contains the early variant exclusively (items 11, 12, 14, 17, 20, 23, 33, 36, 39, 43, 44, 50, 55, 59, 63, 66, 78). For fourteen variables it contains both the early and late variants (items 3, 4, 5, 19, 22, 26, 29, 42, 47, 54, 60, 64, 70, 80). This means that both the number of exclusive late variants and the number of exclusive early variants is higher than in the book of Qoheleth. Figure 17 shows that the concentration of LBH is higher in Ezra than in Qoheleth and that the concentration of EBH is more or less similar. The language of the book of Ezra not only appears later than that of the book of Qoheleth, but it also looks earlier, depending on from which perspective one looks, which means that it is difficult to draw a clear direction in the variation between these two books.

A first inspection of figures 3–16 shows that throughout the selected TBH and LBH books far fewer late variants can be found than their early alternatives. This is not only true for the core LBH books combined, but also for the separate books. Of course each book contains a different set of late variants and early alternatives, but in most cases there are far more early than late tokens. This is confirmed further when all early and late tokens are combined. Figure 17 displays the number of early and late tokens *per 1000*

words of text. It shows that the highest concentration of late tokens can be found in the book of Esther. There the relative amount of the late tokens is about 44%, whereas in the other core LBH books the percentage is between 16% and 30%.

Figures 18–31 show that the number of variables for which there is a complete replacement of an early word or expression by a late variant is low. Most complete replacements can be found in the book of Nehemiah. It contains fifteen variables for which only the late variant is attested, but on the other hand, even here more variables are found for which only the early variants are present (twenty-six items). The book of Nehemiah is followed, in terms of degree of complete replacement, by the books of Esther and Ezra, which contain ten complete replacements each.

One's first impression of figure 31 might be that there are some variables of which an early variant is nearly or completely replaced by a late alternative in the core LBH books of Esther–Chronicles. However the devil is in the details. For example, considering only the eleven variables where there is "complete replacement" (items 21, 25, 27, 28, 32, 34, 37, 58, 72, 73, 79), the spelling of "David" (item 25) is the most credible case,[22] followed by the spelling of "Damascus" (item 27).[23] However, it is hard to consider the single occurrence of זן (item 34) in 2 Chr 16:14 (and also in Ps 144:13) as a "complete replacement" of מִין (in CBH Genesis, Leviticus, and Deuteronomy, and also in Ezek 47:10). Similarly, it is difficult to regard the several occurrences of קרה (item 72) in Neh 2:8; 3:3, 6; and 2 Chr 34:11 (and also in Ps 104:3) as a "complete replacement" of ספן (in CBH Deuteronomy and Kings, and also in Jer 22:14 and Hag 1:4). Furthermore, for seven of the eleven "complete replacements" in Esther–Chronicles the CBH alternatives are very rare in BH (item 58) or occur only once in BH (items 28, 37)[24] or are unknown (items 21, 32, 73, 79).[25]

Figures 14–15 and 29–30 present the results for non-synoptic and synoptic Chronicles. The notable observation is that there are no LBH variants in either figure of synoptic Chronicles, that is, no tokens of the *Lexicon*'s late variants in Chronicles are paralleled in Samuel–Kings.[26] This is not because LBH language is absent from Samuel–Kings or other early writings—far from it[27]—but only

---

[22] However, see Rezetko and Young, *Historical Linguistics*, 456–9.

[23] However, see Rezetko and Naaijer, "Review of Avi Hurvitz *et al.*," §6.2.

[24] Ibid., §5.6.4.

[25] Ibid., §5.6.5.

[26] See n. 12 above on the meanings of synoptic and non-synoptic Chronicles in the present study.

[27] In fact, most typical LBH lexical and grammatical features are attested somewhere in CBH writings. This is illustrated repeatedly in Ian Young, Robert Rezetko, and Martin Ehrensvärd, *Linguistic Dating of Biblical Texts, Volume 1: An Introduction to Approaches and Problems, Volume 2: A Sur-*

because the tokens of the eighty late variants which Hurvitz decided to include in the *Lexicon* are non-synoptic or non-shared between Samuel–Kings and Chronicles. To elaborate further, all of the *Lexicon*'s LBH variants in Chronicles are non-synoptic, that is, all of them are in non-synoptic passages or they are not shared with Samuel–Kings in synoptic passages.

The TBH books contain hardly any of the late variants cited in the *Lexicon*. In general it is assumed that TBH consists of a mixture of early and late variants,[28] but figure 17 shows that there is only a very thin layer of late tokens in Ezekiel, Haggai–Malachi, and the Song of Songs. In the book of Haggai there is a complete absence of late variants; in Malachi and the Song of Songs one can find only one of the late variants in the *Lexicon*; and Zechariah and Ezekiel contain, respectively, five and ten different late variants, generally in very low frequencies.[29]

5.3. DIFFUSION OF EARLY AND LATE VARIANTS IN THE CORPUS OF TBH AND LBH BOOKS

The following figures show the diffusion of early and late variants in the corpus of TBH and LBH books (cf. §4.2). They indicate how widespread the variables used in a specific book are in the other books. Figures 32–42 display the results for each book separately. For instance, figure 32 shows that Ezekiel contains two late variants that occur in only one book, which is Ezekiel, and it contains four late variants that occur in two books, which are Ezekiel and another TBH or LBH book. Figure 43 shows the diffusion of early variants and late alternatives in the eleven selected TBH and LBH books combined.

---

vey of Scholarship, a New Synthesis and a Comprehensive Bibliography (BibleWorld; London: Equinox, 2008); and Rezetko and Young, *Historical Linguistics*. We discussed the example of the spelling of "David" in n. 12 above. Some additional examples are listed in Rezetko and Naaijer, "Review of Avi Hurvitz *et al.*," §6.7.

[28] Hurvitz, *Lexicon*, 3.

[29] This result, to the extent that the late variants in the *Lexicon* are concerned, supports the conclusion of Ehrensvärd and Young, Rezetko, and Ehrensvärd, who argue that the language of Haggai–Malachi is "post-exilic EBH [= CBH]," against the judgment of Shin and Rendsburg. See Martin Ehrensvärd, "Linguistic Dating of Biblical Texts," in Ian Young (ed.), *Biblical Hebrew: Studies in Chronology and Typology* (JSOTSup, 369; London: T&T Clark, 2003), 164–88 (175–87); Young, Rezetko, and Ehrensvärd, *Linguistic Dating of Biblical Texts*, 1:55–6, 106–9, 137–8, etc; Seoung-Yun Shin, "A Lexical Study on the Language of Haggai–Zechariah–Malachi and Its Place in the History of Biblical Hebrew" (Ph.D. diss., Hebrew University, 2007); Gary A. Rendsburg, "Late Biblical Hebrew in the Book of Haggai," in Rebecca Hasselbach and Naʿama Pat-El (eds.), *Language and Nature: Papers Presented to John Huehnergard on the Occasion of His 60th Birthday* (SAOC, 67; Chicago: The Oriental Institute of the University of Chicago, 2012), 329–44.

Figure 32

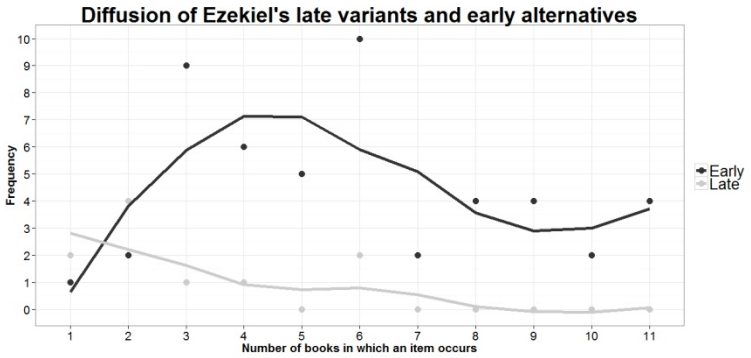


Figure 33

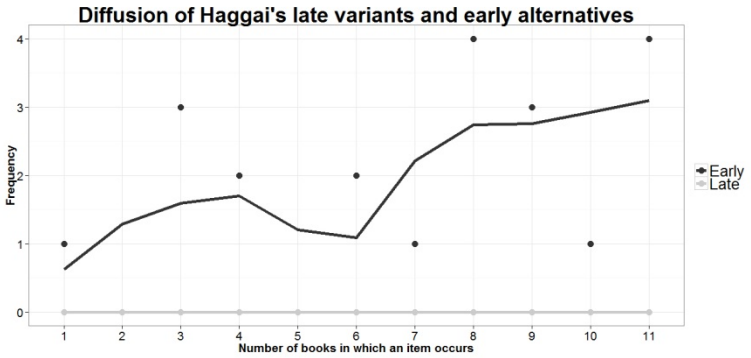


Figure 34

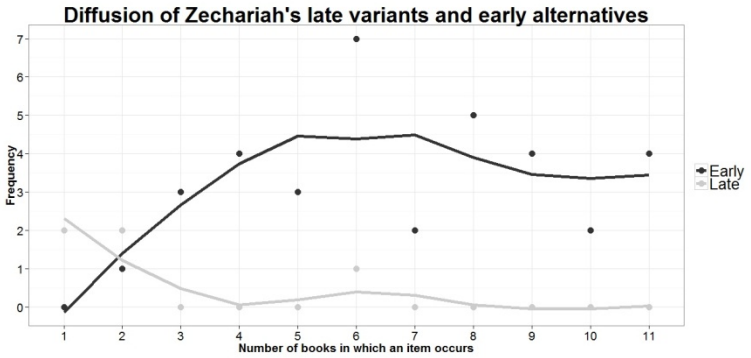


Figure 35

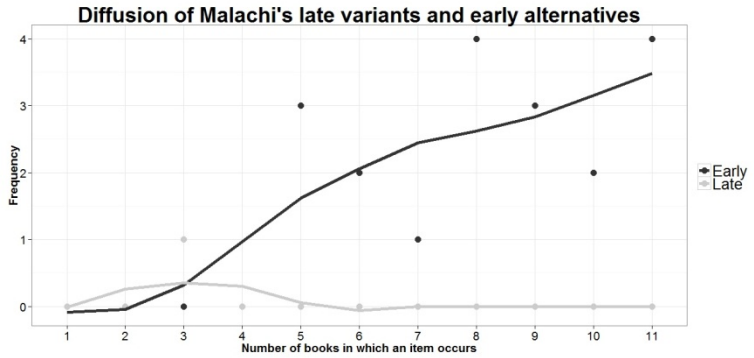


Figure 36

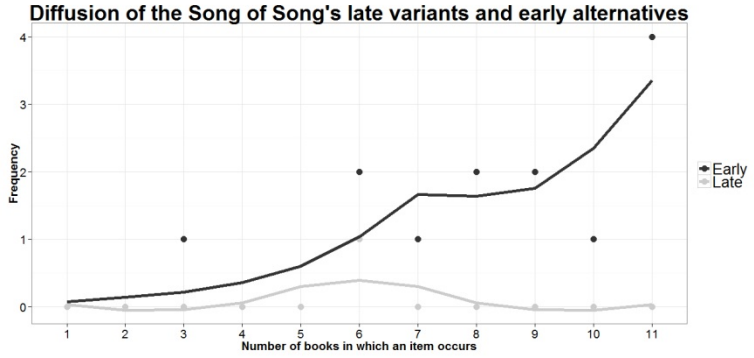


Figure 37

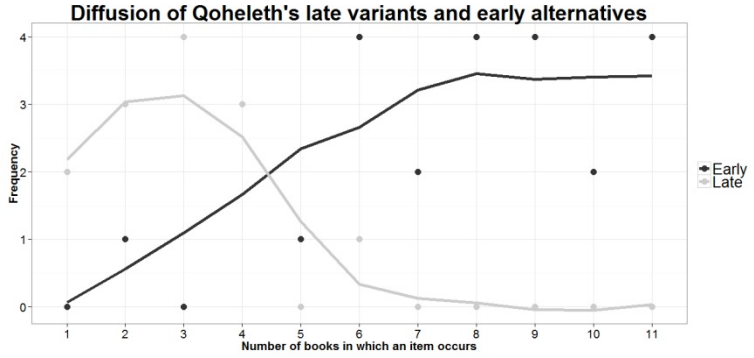


Figure 38

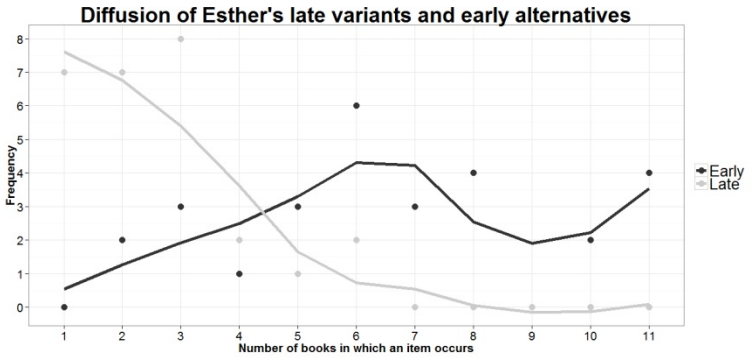


Figure 39

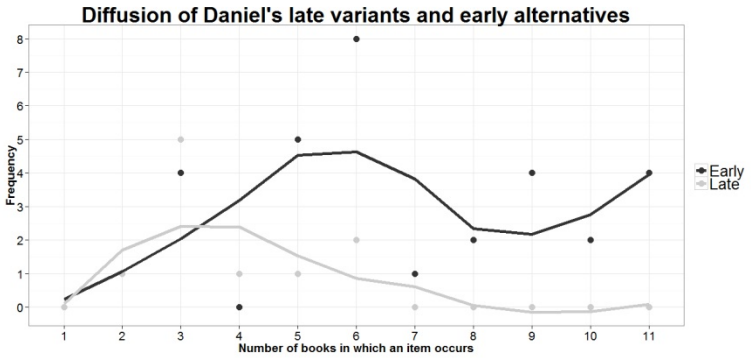


Figure 40

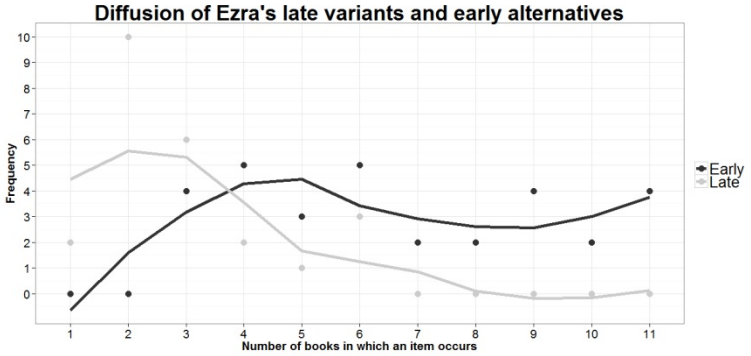


Figure 41

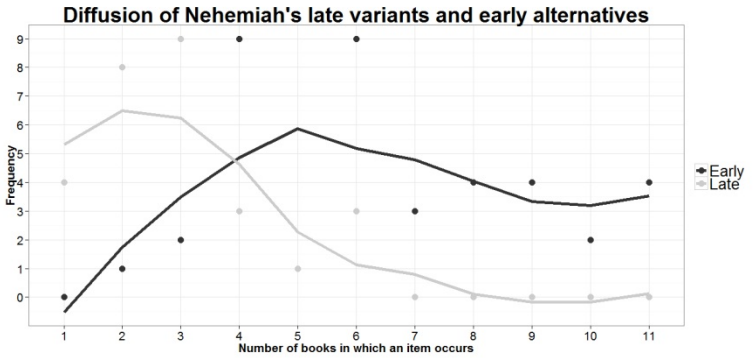


Figure 42

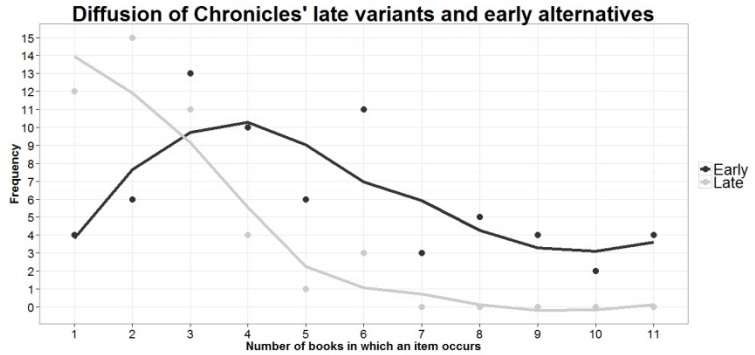
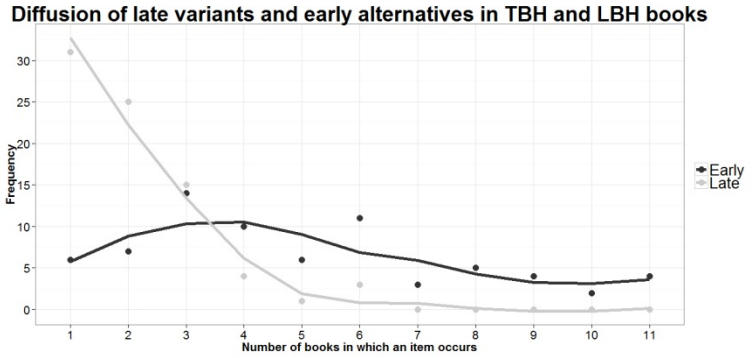


Figure 43

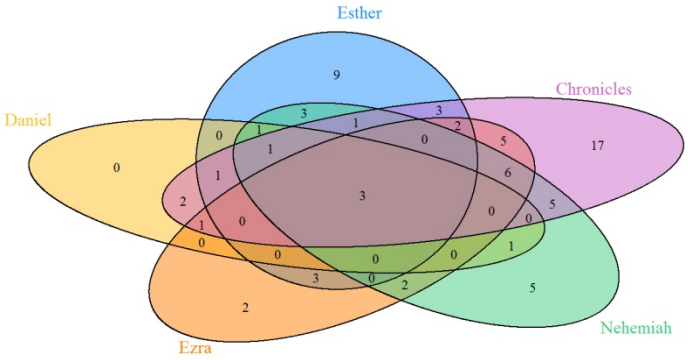


As mentioned above, figure 43 shows the diffusion of early variants and late alternatives in the eleven selected TBH and LBH books combined. The total number of late variants (79) is slightly higher than the number of early alternatives (72), because for some late variants there are no early alternatives in these books (or even

in the entire Bible).[30] Thirty-one of the seventy-nine late variants (39%) can be found in one of the TBH/LBH books only, and if we look at the late variants that occur in three or fewer books (71 of 79, or 90%) it is clear that the rate of diffusion of the late variants in the TBH and LBH books is very low. None of the late variants in the *Lexicon* can be found in more than six of the eleven books under consideration, which makes this impression only stronger. The black smoothing line has no steep segments, meaning that the corresponding early alternatives are spread much more evenly throughout the corpus of TBH and LBH books.

The low rate of diffusion of the late variants can be illustrated in many other ways. For example, the following Venn diagram relates to only the core LBH books of Esther–Chronicles. It illustrates, among other things, that *only three* of the *Lexicon*'s late variants (items 10 [בִּזָּה], 46 [כְּתָב], and 54 [מַלְכוּת]) are shared by all five of these undisputed postexilic books.

Figure 44


Esther
Chronicles
Daniel
Ezra
Nehemiah
9
17
0
2
5
6
3
5
1
2
0
1
0
0
0

guists have dug deeper into Esther's language,[31] and as a result more late variants were found there. For now this remains an open question.

What about the TBH books? Here the numbers of late variants are far lower than in the LBH books, but besides that the curves show strong similarities with those of the LBH books: the early alternatives are more diffused throughout the books under consideration than the late variants. If the diffusion of these late variants and their early alternatives were studied in all the biblical books, we would expect the different rates of diffusion to become even more visible. The rate of diffusion of the late variants remains low, because the CBH books contain late variants in very low concentrations only, but they have early alternatives in a higher concentration than in the books studied here.

What do these observations mean for the linguistic relationships between biblical books and for linguistic dating? In the first place, what is called late or LBH consists mainly of very rare features with a very low rate of diffusion, both in TBH and LBH books. Only a very small part of the late variants can be found in four or more books, and a substantial part is unique for the book in which they occur. In the second place, what the TBH and LBH books share is the high concentration of early alternatives, which are very often distributed widely throughout these books. If one wants to date biblical books of disputed date (that is, books that are not undisputedly late; cf. comments below on linguistic dating generally [§8]) based on the attestation of late variants, it seems necessary that these variants show clear links between the text or book one wants to date and the late corpus as a whole. However, such links turn out to be very weak. Features occurring in only one book do not show the similarities between books; they show only the differences. Therefore, most of the features are unsuitable for dating biblical books. This is not necessarily because these features may not be (relatively or absolutely) late, but because they show that what is generally called late language is a very rare and idiosyncratic part of the language of the corpus of late books.[32]

---

[31] See especially Robert L. Bergey, "The Book of Esther—Its Place in the Linguistic Milieu of Post-Exilic Biblical Hebrew Prose: A Study in Late Biblical Hebrew" (Ph.D. diss., Dropsie College for Hebrew and Cognate Learning, 1983); cf. Martijn Naaijer, "The Common Nouns in the Book of Esther: A New Quantitative Approach to the Linguistic Relationships of Biblical Books" (M.A. thesis, Radboud Universiteit Nijmegen, 2012).

[32] Furthermore, the late language of the undisputed late books is not typical of other potentially late biblical books or postbiblical writings. For discussion of this fact and its implications see Young, Rezetko, and Ehrensvärd, *Linguistic Dating*, *passim*, e.g., 1:276–9; Ian Young, "Late Biblical Hebrew and the Qumran Pesher Habakkuk," *JHS* 8 (2008) (http://www.jhsonline.org/Articles/article_102.pdf).

Nevertheless, we do not think that research on the diffusion of linguistic variants stops here. It has only just begun. In this article we have only studied a restricted number of variables in a restricted number of books. It would be very valuable if similar research were undertaken with more linguistic variables and more biblical books.

5.4. REPLACEMENT VERSUS VARIATION

The results for replacement versus variation (cf. §4.3) are calculated and displayed in two ways, for each type and for each token. *Types* relates to the different things present, and *tokens* to the individual occurrences of something. In our case the types are the eighty variables in the *Lexicon* and the tokens are the individual attestations of those variables. The results can be found in figures 45–46, after which an explanation follows.

Figure 45

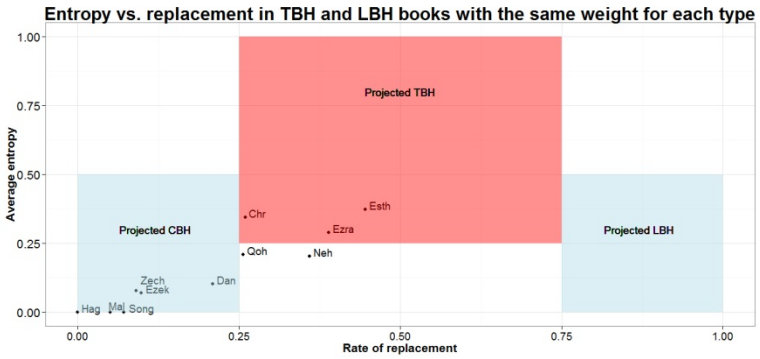
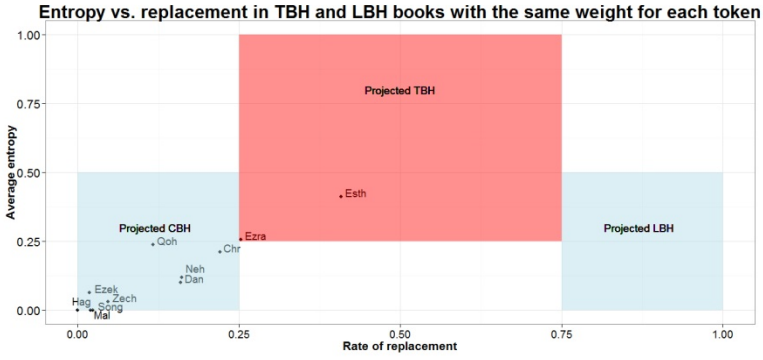


Figure 46



In these figures the average entropy is displayed on the y-axis (vertical) and the rate of replacement on the x-axis (horizontal). The colored rectangles in these figures show the tendency for early variants to be replaced by late variants along the lines of (that is, as projected by) the conventional chronological model of BH. The

block "Projected CBH" shows the expected situation in CBH: the rate of replacement is low and the entropy is low as well. "Projected TBH" shows that replacement is more or less halfway and therefore the amount of chaos cannot be higher, meaning that there is a high entropy. "Projected LBH" shows that the replacement of early variants by late ones is more or less complete, there is a high rate of replacement and a low entropy. Of course the boundaries between these so-called language periods are less sharp than drawn here, so that the colored blocks show a general tendency only.

Some might claim that the "Projected CBH–TBH–LBH" scheme as described in the preceding paragraph does not reflect the reality of language variation and change. In response, first, the scheme essentially reflects in a basic way what many, including Hurvitz in his *Lexicon*, seem to believe about BH: CBH writings generally use early variants, LBH writings generally use late variants, and "transition-period compositions often exhibit intermediate stages of linguistic development, in which the 'old' and 'new' coexist (e.g., the language of Ezekiel . . .)."[33] If "the 'old' and 'new' coexist" in TBH then the logical implication is that they do not "coexist" *as a rule* in CBH or LBH. Second, it is not difficult to compile long lists of examples in well-attested languages where there is a gradual and finally complete replacement of one variant by another one. For instance, samples of Bible translations in Old, Middle, Early Modern, and Modern English illustrate a whole slew of complete replacements—orthographical, phonological, grammatical, and lexical—from earlier to later times.[34] Today, for example, nobody speaks or writes with the subject pronoun *ye* or the third-person singular suffix *-th* (e.g., "hath"), both of which have been completely replaced by *you* and *-s* (e.g., "has"), respectively.[35] Likewise, very many examples of complete replacement are attested in Old, Golden Age, and Modern Spanish. We will give two illustrations. The Arabic loanword for "tailor" in Old Spanish, *alfayate*, was completely replaced by the Catalan loanword *sastre*.[36]

---

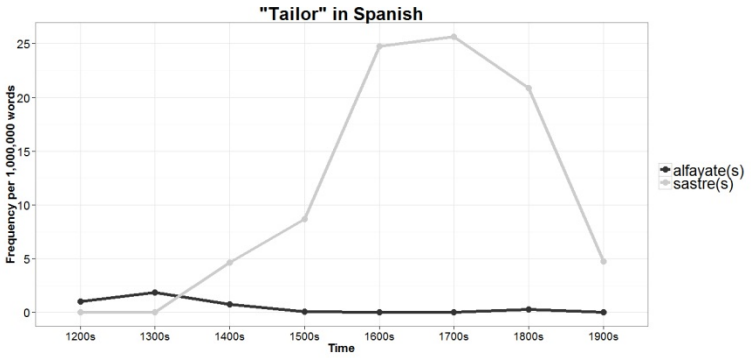
[33] Hurvitz, *Lexicon*, 3.

[34] See, for example, Lyle Campbell, *Historical Linguistics: An Introduction* (3rd ed; Edinburgh: Edinburgh University Press, 2013), 5–8; Hans Henrich Hock, *Principles of Historical Linguistics* (2nd ed; Berlin: de Gruyter, 1991), 2–8.

[35] See Terttu Nevalainen and Helena Raumolin-Brunberg, *Historical Sociolinguistics: Language Change in Tudor England* (Longman Linguistics Library; London: Pearson Education, 2003), 60–1, 67–8, and *passim*.

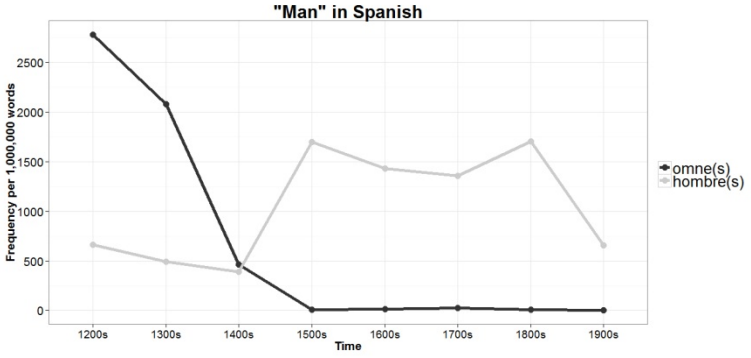
[36] Our database includes all singular and plural forms of these words in the *Corpus del Español* (Mark Davies, *Corpus del Español: 100 million words, 1200s–1900s, 2002–*, available online at http://www.corpusdelespanol.org). For a detailed treatment of the change see Patricia Giménez, "Arabismos en el campo semántico de los oficios: de la competición a la pérdida léxica" (Ph.D. diss., University of Wisconsin-Madison, 2011), 114–41.

Figure 47



Similarly, the predominant word for "man" in Old Spanish, *omne*, was completely replaced by *hombre*.[37]

Figure 48



In short, it is entirely normal among linguists to speak about complete replacements, very often for many variables but at least for some, between the conventional early and late stages of any given language.

In figures 45 and 46 the tendency is more or less similar, but it differs from what might be expected (as explained above). First, the TBH books of Ezekiel, Haggai–Malachi, and also the Song of Songs, show a very low rate of replacement and a low entropy. There is hardly any variation in these books, related to the features in the *Lexicon*. What is more, they show a pattern which might be

[37] Our database includes all singular and plural forms of *omne* and *hombre* in the *Corpus del Español*. We have not included many other less frequent related forms: ome, omes, omme, ommes, home, homes, homine, homines, homne, homnes, ombre, ombres, onbre, onbres, honbre, and honbres. The internal development from Latin to Modern Spanish was: *hominem* > (*h*)*omne* {syncope, *min* > *mn*} > (*h*)*omre* {dissimilation, *mn* > *mr*} > (*h*)*ombre* {epenthesis, *mr* > *mbr*}. For a short summary see Campbell, *Historical Linguistics*, 31.

expected, in the conventional approach, for the early books of the Pentateuch and Former Prophets. Second, the core LBH books of Esther–Chronicles, and also the book of Qoheleth, show a higher rate of replacement than the TBH books, but a higher rate of replacement does not lead to a more consistent use of late variants. The rate of replacement is lower than might be expected in these books and the entropy only increases with a higher rate of replacement. These books show the pattern which might be expected for TBH books.

In the conventional chronological model of BH it is expected that there is a gradual replacement of early variants by late ones. This replacement does not seem to have taken place in reality, based on the items in the *Lexicon*. There is only an increase of variation. This does not mean that the history of BH is not visible in these data. It is possible that they reflect diachronic change, but the traditional model based on complete, gradual replacement, such as with the S-curve, is insufficient to describe them properly. Furthermore, as discussed above (§5.2), complete replacement is uncommon or problematic for individual variables and it is unattested when multiple variables are examined simultaneously; in short, different books have different features with strongly varying properties. A mix of other explanations may give more insight into the mechanisms of variation and change in BH, related to models that are based on actual patterns found in the data, instead of patterns that are presupposed by certain theories. Therefore it would be interesting to calculate the entropy of the separate non-biblical DSS and Mishnah tractates, based on the items in the *Lexicon*.

5.5. TAKING STOCK OF SOME RESULTS UP TO THIS POINT

In our study of TBH and LBH we have followed a different approach than is usual in studies of BH. In general one book or text is selected, after which the number of late variants in it is counted. We started with eighty late variants and counted these and their early alternatives in eleven TBH and LBH books to get an impression of linguistic variation within and between these books. Of course one may find other results if other features are studied, but we do not think that those results would compromise our conclusions, because it may be assumed that the features in the *Lexicon* are good representatives of "Late Biblical Hebrew" in general.[38]

In the core LBH books of Esther–Chronicles only a relatively thin layer of late language can be found. In all of these books there are items with rates of replacement varying between 0% and 100%, but in all of them there are (far) more variables for which there is an exclusive use of the early variant than an exclusive use of the late one. If the total number of early and late tokens is compared in each of the core LBH books we see that Esther is the latest book

[38] Rezetko and Naaijer, "Review of Avi Hurvitz *et al.*," §5.7.1.

with about 44% late attestations. The other core LBH books have lower concentrations of late language, between 16% and 30%.

There are only weak linguistic links between the TBH books and the LBH books, at least concerning the late variants in the *Lexicon*. Only a very thin layer of late language can be found in the TBH books of Ezekiel, Haggai, Zechariah, and Malachi, and also in the Song of Songs. However, strong links are evident between all these books (and therefore also with CBH books) concerning early variants, both in the number of early variants found in these books and the rate of their diffusion throughout the corpus of eleven TBH and LBH books.

Not only do the TBH books have weak links with the LBH books in regard to the late variants, but the links between the individual core LBH books are also weak. About 90% of the late variants are attested in only three or fewer TBH/LBH books, while the rate of diffusion of their early alternatives is generally higher. *It can be concluded that most late variants are very rare and idiosyncratic.*

In the conventional diachronic model, and especially in those cases in which the S-curve is used, it is often assumed that linguistic change takes place in a certain direction. It is assumed that in the preexilic period there was CBH writing, which was replaced gradually by LBH writing in the postexilic period. As a minimum it is assumed that at least some early variants in CBH are replaced by some late alternatives in LBH. However, our analysis has shown that complete replacement hardly ever took place, and in fact it is generally ill-advised to speak about replacement in any meaningful sense of the word. In all the core LBH books the amount of early language is higher than the amount of late language. However, it is possible to show a certain direction. Books with a higher rate of replacement have a higher internal variety of relatively rare variants. The direction of change is not one of early variants to late variants, but from early variants to early *and* late variants. We believe that those who study linguistic variation in BH should start searching for models in addition to the model of replacement to describe this variation. For example, diffusion and entropy could be useful notions to keep in mind for doing this.

6. REMARKS ON "CHANGE" AND "CONTINUITY," OR HURVITZ'S "REPLACEMENT" AND "COEXISTENCE"

At various points the *Lexicon* refers explicitly to "change" in BH[39] and elsewhere the idea of change from CBH to LBH is implicit. For example, in the discussion of the *piel* of קבל,[40] Kutscher is cited in support of the idea that there was a "change" (twice) in usage from לקח in early BH to קבל in late BH.[41] Then, a few sentences

[39] Hurvitz, *Lexicon*, 1, 3, 35, 82, 186, 215, 229, 232, 233, 235.
[40] Ibid., 213–6.
[41] Ibid., 215.

later, there is a citation from Bergey where he points out that לקח is the preferred verb in LBH and the DSS.[42] This raises again the thorny issue of "change" and "continuity" in BH or, as discussed above, "replacement" and "coexistence."[43]

This brings us to a key theoretical-methodological problem in historical linguistics: the unfortunate focus on change to the detriment of continuity, the latter constituting the lion's share of any language through time. Already Ferdinand de Saussure emphasized this: "What predominates in all change is the persistence of the old substance; disregard for the past is only relative. That is why the principle of change is based on the principle of continuity."[44] In reality, much of what language scholars refer to as "change" is not change in the sense of "replacement" but "instances of change in progress on the basis of different rates of variation"[45] or change in "the markedness features" of the variables studied.[46]

Returning to the *Lexicon*, our presentation and discussion of data above (§4.3 and §5.4) have shown that continuity of early language in late (TBH and LBH) writings is the norm. In fact, and on the whole, early language predominates in those writings, such that we are justified to speak about large-scale "invariability" or "invariation" of early language throughout all of BH. Furthermore, the

---

[42] Ibid., 215. For a discussion of these verbs from the perspective of a variationist approach see Rezetko and Young, *Historical Linguistics*, 304–7.

[43] Rezetko and Naaijer, "Review of Avi Hurvitz *et al.*," §5.7.5.3 and §5.7.5.4.

[44] Ferdinand de Saussure, *Course in General Linguistics* (trans. Wade Baskin; ed. Perry Meisel and Haun Saussy; New York: Columbia University Press, 2011), 74.

[45] "A very dangerous line of argumentation seems to have developed within sociolinguistics . . . with respect to the nature of variation and change. For some, the implicational relationship between the two holds equally well in both directions so that not only does change imply variation, but also variation implies change. Most sociolinguistic studies in recent years are looking for, and usually report, instances of change in progress on the basis of different rates of variation for age groups, social classes, etc., while relatively few studies have devoted themselves to the analysis of situations in which change does not occur" (Suzanne Romaine, *Socio-Historical Linguistics: Its Status and Methodology* [Cambridge Studies in Linguistics, 34; Cambridge: Cambridge University Press, 1982], 200).

[46] "Linguistic innovations and older forms frequently live side by side, often so that the new form is marked and the old unmarked. As time passes, and the new form slowly becomes better and better established in the language, it gains more ground, and the markedness features change places. The new form becomes the unmarked one, and the old form is confined to special uses until it is perhaps dropped out of the language completely" (Helena Raumolin-Brunberg, *The Noun Phrase in Early Sixteenth-Century English: A Study Based on Sir Thomas More's Writings* [Mémoires de la Société Néophilologique de Helsinki, 50; Helsinki: Société Néophilologique, 1991], 23).

notion of "coexistence" of early and late linguistic variants is more appropriately applied to the LBH books than to the TBH ones. This may come as a surprise to some scholars of BH. In addition, the *Lexicon*'s assertions about "change" and "replacement," or the "living" language and "survivals," are demonstrated to be problematic at best and erroneous at worst. How, for example, are the "old CBH counterparts" of the verb בעס "nothing more than ancient survivals or archaizing devices"[47] when those "old CBH counterparts" occur exclusively in the late books of Zechariah, Song of Songs, Esther, and Ezra, and the early and late variants appear together in the late books of Qoheleth, Nehemiah, and Chronicles? In our opinion it is actually rather dubious whether there is any change as such, in the sense of completed change or replacement, attested in BH as a whole, and there is certainly no such change in connection with any late variants that occur *often* in late writings. From a different perspective, common phenomena like *ye/you* and *-th/-s* in the history of English and *alfayate/sastre* and *omne/hombre* in the history of Spanish, are not paralleled by similar examples of complete replacement between the early and late stages of BH.[48] That does not mean that some individual books do not attest the replacement of early variant A by late variant B (e.g., יֵשׁוּעַ for יְהוֹשׁוּעַ in Ezra and Nehemiah; contrast Haggai, Zechariah, and Chronicles), but as soon as the focus moves from individual books to sets of books, especially books that are conventionally dated to the same historical period or language state, exceptions or discrepancies inevitably appear.[49]

7. REMARKS ON PERIODIZATION, STATES AND TRANSITIONS

Hurvitz's linguistic dating method is inseparable from the assumption of distinct language periods, transitional periods between those periods, and books acknowledged to have been written (more or less as they are found in the MT) in one or another of those periods.[50] Elsewhere we have examined various theoretical and methodological problems with the notion of language periodization.[51]

---

[47] Hurvitz, *Lexicon*, 143–4.

[48] This is quite surprising given that around one-thousand years presumably separated the original composition of the earliest and latest BH writings.

[49] This problem is compounded when literary and textual developments *within* individual biblical books are taken into consideration.

[50] Rezetko and Naaijer, "Review of Avi Hurvitz *et al.*," §5.5.

[51] Young, Rezetko, and Ehrensvärd, *Linguistic Dating*, 1:49–58; Rezetko and Young, *Historical Linguistics*, 49–56, 395–402. In short, "we believe together with some other Hebraists that the conventional three-stage model of Biblical Hebrew—Early Biblical Hebrew and Late Biblical Hebrew with a transition between them—is problematic in a variationist framework, offers no workable basis for empirical research, and should be

The point we want to emphasize here is that if CBH, TBH, and LBH are to be regarded as roughly three consecutive phases of language, then the actual distribution of the linguistic data (that is, the data in the *Lexicon*) in the eleven books we have studied (Ezekiel, Haggai–Malachi, Song of Songs, Qoheleth, Esther–Chronicles) argues in favor of the view that these books represent the language of (early) CBH or TBH, but not LBH (cf. figures 45–46). We suggest that some explanations that could help to shed light on the distribution in these books of linguistic variables, especially the sporadic appearance of late variants, include, for example, distinct editorial and scribal processes,[52] or the possibility that the writers who had occasional recourse to the late variants were *early (and only partial) adopters* of new linguistic variants.[53]

8. REMARKS ON HURVITZ'S "LINGUISTIC DATING"

In our separate review we show that the ultimate purpose of the *Lexicon* is to reinforce the *linguistic dating* of early *writings* to the early period and late *writings* to the late period.[54] We have argued elsewhere that rates of (non-)accumulation of late variants in various (late) biblical and postbiblical writings,[55] the fact that there can be early, middle, and late adopters of linguistic innovations,[56] and the general conclusion reached by literary and textual critics that biblical writings—including those written in CBH—did not stop developing until late in the Second Temple period, all problematize Hurvitz's linguistic dating approach. In this article we have pro-

---

set aside in favor of less idealized and more rigorous descriptive approaches to the database" (Rezetko and Young, *Historical Linguistics*, 408). Furthermore, the suggestion that there were transitions from BH to QH and/or RH is very problematic (Young, Rezetko, and Ehrensvärd, *Linguistic Dating*, 1:197, 241–3, 246–8, 277–8; cf. additional literature cited there).

[52] Or, loosely spoken, "style" (Rezetko and Young, *Historical Linguistics*, 400–2, 599; cf. cross-references and additional literature cited there).

[53] For discussion of "early adopters" and other language users in the variationist sociolinguistic framework see Rezetko and Young, *Historical Linguistics*, 223–8; cf. 318–27, 395–402, 407. With reference to the subtitle of the *Lexicon, Linguistic Innovations in the Writings of the Second Temple Period*, it seems unlikely that the late variants are "innovations" by the writers of the books in which they occur. James Milroy observes that "there seems to be no easy way for empirical studies of change in progress to identify in the data the crucial distinction between innovators and early adopters," and, "the individuals or groups that we identify as carrying linguistic changes are likely to be early adopters of the change rather than innovators" (James Milroy, *Linguistic Variation and Change: On the Historical Sociolinguistics of English* [Language in Society, 19; Oxford: Blackwell, 1992], 184, 201).

[54] Rezetko and Naaijer, "Review of Avi Hurvitz *et al.*," §3.

[55] Young, Rezetko, and Ehrensvärd, *Linguistic Dating*, *passim*.

[56] Rezetko and Young, *Historical Linguistics*, *passim*.

vided additional criticism of Hurvitz's approach by showing that what is generally called late language is a very rare and idiosyncratic part of the language of the corpus of late books and that these books show very different tendencies in their use (including non-use) of late variants.[57] In our opinion it is time to set aside linguistic dating as a viable objective of the historical linguistic study of BH in order to focus our efforts on more thorough descriptive approaches to the language of ancient Hebrew.

9. CONCLUSIONS

First, our main objective has been to present an alternative, quantitative approach to describing the lexical stock of late biblical writings.[58] We took as our starting point the late and early variants cited in Hurvitz's *Lexicon* and sought to document and visualize their distributional patterns in a selection of TBH and LBH writings. We must emphasize that our aim in this article was not to contest that the late variants cited in the *Lexicon* could actually be late, relatively or absolutely.[59]

Second, our main conclusion is that the late language of late biblical writings is rare and idiosyncratic. Late language is uncommon in comparison with the high incidence (or continuity) of early language in the late corpus. Furthermore, late language is not dif-

---

[57] Consequently, it is arguable, and in fact it has been argued, that late writers and editors were fully capable of selecting regularly either late or early variants. Accordingly, whereas the presence of late variants might prove lateness, the absence of late variants would not prove earliness. In addition to the volumes cited in the preceding footnotes see, for example, Robert Rezetko, "The Qumran Scrolls of the Book of Judges: Literary Formation, Textual Criticism, and Historical Linguistics," *JHS* 13 (2013) (http://www.jhsonline.org/Articles/article_182.pdf); idem, "The (Dis)Connection between Textual and Linguistic Developments in the Book of Jeremiah: Hebrew Bible Textual Criticism Challenges Biblical Hebrew Historical Linguistics," in Raymond F. Person, Jr. and Robert Rezetko (eds.), *Empirical Models Challenging Biblical Criticism* (SBLAIL 25; Atlanta: SBL Press, in press).

[58] The method we have employed is not subject to the criticism of "counting" or "numbering" as some Hebraists might want to claim, citing for example Driver's criticism of Giesebrecht (e.g., Avi Hurvitz, "The 'Linguistic Dating of Biblical Texts': Comments on Methodological Guidelines and Philological Procedures," in Cynthia Miller-Naudé and Ziony Zevit [eds.], *Diachrony in Biblical Hebrew* [LSAWS, 8; Winona Lake, IN: Eisenbrauns, 2012], 265–79 [274]). First, the variationist approach we have used was developed since Driver's days and is widely used in historical studies of other languages. Second, "counting" words and other linguistic elements helps us to move beyond intuitional judgments, all too common in studies of BH, by drawing boundaries that allow a more objective weighing of data.

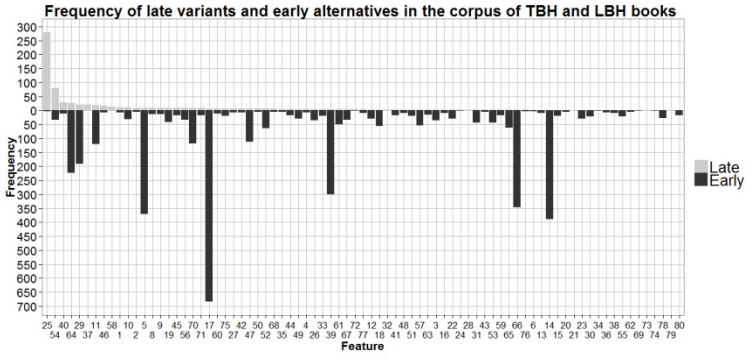
[59] However, see Rezetko and Naaijer, "Review of Avi Hurvitz *et al.*," §6.7.

fused throughout the individual writings in the corpus. Finally, complete replacements of early variants by late ones are scarce for individual books and unattested for the corpus as a whole. LBH is very heterogeneous.

Third, the preceding results problematize linguistic periodization and linguistic dating. Specifically, talking about a corpus of late writings as representative of a state of linguistic development is highly problematic.

Fourth, and finally, to return to one of our questions at the start of this article, "What *is late* about the lexicon of LBH?," our answer is, "not much."[60] The following figure displays all the variables given in the *Lexicon* according to the descending frequency of late variants in all the late books we have examined here (Ezekiel, Haggai–Malachi, Song of Songs, Qoheleth, Esther–Chronicles).[61]

Figure 49



Once again, this final figure underscores how rare and idiosyncratic late language is in these late books combined. In summary, the lexicon of "Late Biblical Hebrew" looks far more "early" than "late."[62]

---

[60] The corpus of "Late Biblical literature" in the *Lexicon* includes Second Isaiah, Haggai–Malachi, Qoheleth, and Esther–Chronicles (Hurvitz, *Lexicon*, 4). Our study has included nine of these ten books, excluding Second Isaiah (which attests only two of the *Lexicon*'s late variants, items 42 [one token!] and 56 [three tokens!]; cf. Ian Young, "'Loose' Language in 1QIsa[a]," in Shani Tzoref and Ian Young [eds.], *Keter Shem Tov: Essays on the Dead Sea Scrolls in Memory of Alan Crown* [PHSC, 20; Piscataway, NJ: Gorgias, 2013], 89–112 [esp. 98–100]). Instead, we included Ezekiel and the Song of Songs in our study.

[61] Figure 16 pertains only to the core LBH books of Esther–Chronicles.

[62] As the present article was going to press we encountered Steven E. Fassberg, "What is Late Biblical Hebrew?," *ZAW* 128 (2016): 1–15. The article is intended as a basic summary of research on LBH and its linguistic characteristics. Unfortunately, the article reverberates varied shortcomings of much other contemporary work on LBH. As one example of many, in the "list of some of the most salient features of Late Biblical Hebrew" (ibid., 11–4), Fassberg cites (p. 14) the noun תַּעֲנִית ("fast") in

Ezra 9:5 as a ("salient") feature of LBH in contrast with CBH צוֹם in passages like "2 Sam 12:16–20" (*sic*; 2 Sam 12:16, 20–23?). (תַּעֲנִית is item 80 in the *Lexicon*.) However, if the question were "What is the 'salient' word for 'fast' in LBH?," undoubtedly the answer would be צוֹם, not תַּעֲנִית, since the latter occurs only one time in (L)BH, in Ezra 9:5, whereas the "early" item occurs exclusively and frequently in most late biblical writings, that is, *it* is the "salient" LBH feature (Isa 58 [= III Isaiah]:3 [x2], 4 [x2], 5 [x2], 6; Zech 7:5 [x3]; 8:19 [x4]; Esth 4:3, 16 [x2]; 9:31; Dan 9:3; Ezra 8:21, 23; Neh 1:4; 9:1; 1 Chr 10:12; 20:3 [non-synoptic]; cf. Jer 14:12; 36:6, 9; Joel 1:14; 2:12, 15; Jonah 3:5; Ps 109:24). In short, Fassberg's article is not really very helpful in answering the question "What is Late Biblical Hebrew?"