Time Adverbial and Verb Collocations in Biblical Hebrew

Cody Kingham

PhD candidate, Hebrew Studies

St John's College, University of Cambridge

24 April 2019

**Introduction**

One of the most productive and interesting areas of Biblical Hebrew grammar is that of time expression, as shown by the continual stream of research on the Biblical Hebrew verbal system.[[1]](#footnote-1) Yet, considerably little attention has been paid to time adverbials.[[2]](#footnote-2) This despite the fact that time adverbials are known to serve important roles in establishing time reference in language.[[3]](#footnote-3) In Biblical Hebrew, time adverbials can anchor discourse,[[4]](#footnote-4) modify event time or aspect,[[5]](#footnote-5) and direct pragmatic focus.[[6]](#footnote-6) While time adverbials are mentioned in most major grammars, there is as of yet no comprehensive treatment on their composition, statistical distribution, and function.[[7]](#footnote-7) Grammars tend to treat time adverbials as a series of isolated features, divided amongst adverbs, adverbial accusatives, or prepositions.[[8]](#footnote-8) This approach leads to an atomistic, word-centric focus, and fails to explain how the various components mutually interact and together function as sentence arguments.

This study addresses these shortcomings by treating time adverbials as adjuncts composed of entrenched, constructional patterns.[[9]](#footnote-9) Constructions blend and order patterns of morphemes, lexemes, phrases, and sentence types to convey meaning.[[10]](#footnote-10) The patterns can be identified statistically by calculating collocation frequencies between the units. Associated forms co-occur frequently while disassociated forms avoid co-occurrence.[[11]](#footnote-11) Combined with techniques from computational linguistics and Construction Grammar, these patterns make a quantitative approach to semantic analysis possible.[[12]](#footnote-12) The full project aims to provide a comprehensive inventory of time adverbial constructions in Biblical Hebrew along with their sentential and textual functions. This brief essay provides a preliminary sampling of that effort.

Hebrew grammars typically divide time adverbials into three separate rubrics of adverbs, adverbial accusatives, and prepositions.[[13]](#footnote-13) The adverbs include particles like עַתָּה, טֶרֶם, שִׁלְשׁוֹם, מָחָר, and עוֹד.[[14]](#footnote-14) Substantives which function adverbially are considered under "adverbial accusatives."[[15]](#footnote-15) Though historically related to early Northwest Semitic cases,[[16]](#footnote-16) the Biblical Hebrew accusative is now only marked syntactically.[[17]](#footnote-17) It is an "optional" adjunct to the verb which indicates time, location, manner, or other specifications.[[18]](#footnote-18) Terms attributed to the time accusative include יוֹם, בֹּקֶר, צָהֳרַיִם, and שָׁנָה, which can appear in larger noun or prepositional phrases.[[19]](#footnote-19) The grammars typically ascribe no more than two senses to time substantives: 1. points in time ("when?") and 2. durations ("how long?").[[20]](#footnote-20) Information about time positioning is nested separately under the individual prepositions בְּ, לְ, כְּ, אַחַר/אַחֲרֵי, עַד, בַּיִן, לִפְנֵי, and מִקֵץ.[[21]](#footnote-21) Prepositions orient an event ("trajector") in relation to a time noun ("landmark").[[22]](#footnote-22) Temporal prepositions are considered metaphorical extensions from the spatial sphere.[[23]](#footnote-23)

There are only two studies dedicated to time adverbials in Biblical Hebrew, those of DeVries (1975) and Brin (2001).[[24]](#footnote-24) DeVries's study is dedicated to יוֹם as "the basic unit of time" in "Hebrew thought."[[25]](#footnote-25) He discerns two main constructions for conveying past, present, and future: בַּיוֹם הַהוּא (past/future) and הַיוֹם / הַיוֹם הַזֶה (present).[[26]](#footnote-26) For each of the three formulas DeVries reviews the primary literary contexts, attending to source critical concerns.[[27]](#footnote-27) Brin's later work is more expansive. He recognizes that time adverbials depend on reference points, typically a person, event, or period known from the text or history.[[28]](#footnote-28) Functionally, the points locate events later, simultaneous, or preceding them.[[29]](#footnote-29) He identifies יוֹם as the quintessential time, with 2,317 occurrences.[[30]](#footnote-30) X-יְמֵי (plural) indicates a duration relative to X, while X-יוֹם (singular) indicates a point relative to X.[[31]](#footnote-31) Brin treats an assorted variety of time markers. דוֹר indicates "pieces of history"; לְפָנִים and בְּרִאשׁוֹנָה are distant past.[[32]](#footnote-32) Under "durations" Brin lists מִן...וְעַד,מַעְלָה /מִן...וְהָלְאָה, and עַד הַיוֹם הַזֶה.[[33]](#footnote-33) He groups רֶגַע, קָרוֹב, מָחָר, תְּמוֹל, לָיְלָה, and יוֹם as "brief" units and לֹא קָרוֹב, לְעִתִּים, אַחֲרוֹן, רָחוֹק, לְעוֹלָם, and לְעַד as "long periods."[[34]](#footnote-34) He notes the primary calendrical units יוֹם, שְׁבוּעַ, חֹדֶשׁ, שָׁנָה and the seasons זֶרַע וְקָצִיר, קוֹר וְחוּם, קַיִץ וְחֹרֶף.[[35]](#footnote-35) Under expressions for past Brin lists לְרִאשֹׁנָה, לְפָנִים, מֵעוֹלָם, בְּטֶרֶם, עוֹלָם מִימֵי, יְמֵי קֶדֶם, and דוֹר ודוֹר.[[36]](#footnote-36) For present are הַיוֹם, בָּעֵת הַזֹּאת, הָאֵלֶּה בַּיָּמִים, עַתָּה, הַפַּעַם, and עַד הַיוֹם הַזֶה.[[37]](#footnote-37) Finally, for "future" he gives מָחָר andהַדּוֹר הָאַחֲרוֹן.[[38]](#footnote-38)

DeVries and Brin primarily examine historical-exegetical issues and not linguistic concerns. In a 1997 article, Van der Merwe recognizes the lack of such data, and calls for an approach informed by studies of time adverbials in linguistics.[[39]](#footnote-39) For instance, Quirk identifies six different times in English: 1. time position, 2. forward span, 3. backward span, 4. duration, 5. frequency, and 6. relationships (e.g. "up to that time").[[40]](#footnote-40) Harkness distinguishes between "anchor-dependent" (e.g. "last year," speech-time dependent) and "anchor independent" (e.g. "in 1978").[[41]](#footnote-41) Van der Merwe outlines several questions yet to be answered in Biblical Hebrew, including whether certain syntactic features in time adverbials correlate with duration, frequency, and position meanings, whether syntactic classes correlate with dependent/independent anchors, whether position in the sentence affects time semantics, whether there are differences between temporal and non-temporal adjuncts, and whether there are any other semantic classes unique to Biblical Hebrew.[[42]](#footnote-42) Answering these questions requires the "compilation of a taxonomy of BH temporal adjuncts."[[43]](#footnote-43) He also argues for treating time "adverbials," which are phrases that indicate points in time as a sentence constituent ("adjunct"), rather than word-class categories.[[44]](#footnote-44) These questions are the impetus for the present study.

Addressing these diverse concerns requires a new linguistic framework. Traditional grammar divides language into distinct regimes of syntax, semantics, and pragmatics, as well as morphemes, words, phrases, and sentences.[[45]](#footnote-45) It is for this reason that information about time adverbials is scattered across disparate sections of the grammars. Yet time adverbials, as do many forms, show little regard for these boundaries.[[46]](#footnote-46) Particles such as עַתָּה and אָז, for instance, regularly shift from a temporal sense to a pragmatic sense, depending on the syntactic pattern enclosing them.[[47]](#footnote-47) Or, another example, the demonstrative force of הַ in הַיוֹם is limited to time adverbials.[[48]](#footnote-48) Likewise, words appear to ignore word classes. When יוֹם is used without any other modifiers, it functions like a bare adverb (e.g. \*); vice versa, when עוֹלָם is used with a plural noun ending, it functions like a noun (e.g. \*). These phenomena suggest that syntactic, semantic, and pragmatic meaning derive from idiomatic patterns rather than word-inherent classes.

The key insight of Construction Grammar is that there are no word-inherent, universal grammatical categories.[[49]](#footnote-49) Rather, individual languages consist of unique pattern-to-meaning pairings. Every unit in a language, from morphemes to sentence arguments, consists of fixed or schematic patterns.[[50]](#footnote-50) An "adverb" is therefore not an inherent word class, but a pattern of use which a given word can become associated with.[[51]](#footnote-51) These patterns, like lexemes, carry syntactic, semantic, and pragmatic meaning. Thus Construction Grammar makes no sharp distinction between the three modules.[[52]](#footnote-52) The boundary between grammar and lexicon is likewise blurred as forms are shown to be inter-dependent and associated.[[53]](#footnote-53) The table below illustrates the lexeme-syntax continuum in Biblical Hebrew time adverbials.

# Figure : Continuum of Fixity and Schematicity between Time Adverbials in Biblical Hebrew[[54]](#footnote-54)

fixed

fixed

schematic

|  |  |
| --- | --- |
| morpheme | ִים (pl), ַיִם (du), ָם / וֹם (adv)[[55]](#footnote-55) |
| word | מָחָר, לָיְלָה, עַתָּה |
| complex word | מִקֵץ, לְעוֹלָם, לִפְנֵי, וְעַתָּה |
| complex word (partially filled) | יָמִים, np patterns |
| idiom (filled) | בֵּין הָעַרְבֳּיִם |
| idiom (partially filled) | בֶּן + cardinal + time; (age)  time + time (e.g. יוֹם יוֹם, "day by day") |
| attributive time NP | הַ + time + הַ + attr |
| macrosyntactical indicator | וַיְּהִי + time |
| durative aspectual | ø + quantification + time + verb |

Constructions at the top of the table are more fixed whereas those on the bottom have more complex, schematic parts. Constructions function idiomatically: the ִים morpheme is paired with plural meaning; מִקֵּץ has lost its usual sense of "detachment" and is appended to durations;[[56]](#footnote-56) the plural יָמִים conveys a duration rather than merely "plural" days;[[57]](#footnote-57) וַיְּהִי + time conveys pragmatic meaning, orienting the reader by setting the narrative time.[[58]](#footnote-58) These examples, and others, exhibit the same unpredictability and idiomaticity attributed to the lexicon.[[59]](#footnote-59) Construction grammarians aim to identify and inventory the forms and functions of such idiomatic units, called "constructions."

This study uses the tools of collocation analysis, namely, statistical association measures and computer algorithms, to identify time constructions. This approach is grounded on a usage-based view of language, whereby constructions obtain meaning through their habitual use in a given context.[[60]](#footnote-60) The more often a given construction is used in a certain way, the more entrenched the connection becomes in a user's mind.[[61]](#footnote-61) Constructions, as do words, frequently collocate with semantically informative patterns.[[62]](#footnote-62) For example, the plural morpheme construction and cardinal numbers co-occur frequently in Biblical Hebrew due to their interdependent meaning.[[63]](#footnote-63) Dependency can be operationalized with a statistical measure of association. The measure answers if, given two constructions A and B, their co-occurrence count X, and their respective co-occurrences with all other constructions in the corpus, it is likely that A and B are interdependent.[[64]](#footnote-64) This method is currently being utilized in lieu of grammaticality judgments to test linguistic theories in cognitive linguistics and Construction Grammar.[[65]](#footnote-65) It is therefore especially well-suited for a language without living informants. Furthermore, this approach enables an investigation into the interdependencies between the various components of Hebrew time adverbials.

A collocational investigation of time adverbials in Biblical Hebrew requires broad access to corpus data. The *BHSA* of the Eep Talstra Centre for Bible and Computer is a scholarly, open source syntactic database that contains labels on phrases with time adverbial function.[[66]](#footnote-66) The database covers the whole Hebrew Bible. The time data has gone through a preliminary manual check for accuracy for this project.[[67]](#footnote-67) While it is possible *BHSA* misses some limited cases, the source provides a broad and diverse initial dataset from which primary tendencies can be learned. Text data is processed and counted in Python using a corpus analysis tool, Text-Fabric, alongside statistics packages.[[68]](#footnote-68) The statistical method of Collostruction Analysis is used to detect the associations.[[69]](#footnote-69)

For this analysis "time adverbial" shall refer to any phrasal, sentence constituent which contributes time reference to a verbal event structure. This pilot study excludes adverbials headed by the more pragmatic particles אָז, עַתָּה, כֵּן, and אַךְ. Linguistic concepts will be represented in small caps. The analysis has four primary goals: 1) to catalogue and count the distribution and diversity of time adverbials, 2) to develop a data-driven semantic taxonomy, 3) to measure positional tendencies within sentences, 4) to measure collocational tendencies with verbal aspect classes ("*Aktionsart*"). Theoretical guidance is provided by Haspelmath's study of time adverbials in world languages, Fillmore's constructional analysis of time adverbials, and Croft's analysis of verbal event frames.[[70]](#footnote-70)

**Time Adverbials in Biblical Hebrew**

The analysis of time adverbials in Biblical Hebrew begins with an exploratory overview of the identity and distribution of phrases marked for adverbial time in the *BHSA*. The *BHSA* divides the Hebrew text into word, phrase atom, phrase, clause atom, clause, and sentence objects.[[71]](#footnote-71) The phrase object is the largest functional unit without predication and clauses are the most basic unit with predication. Every phrase in the database has a feature called "function," which describes its argument role in its enclosing clause. A total of 29 functions are encoded, including predicate (i.e. main verb), subject, object, complement, time, location, and adjunct (where adjunct contains sub-functions which are neither time nor location).[[72]](#footnote-72) *BHSA* phrases also have a "type" (typ) feature which tells its formal type, i.e. np ("noun phrase" etc.), pp, advp.

*BHSA* is accessed and processed in Python. The data is available through a Python package called Text-Fabric.[[73]](#footnote-73) Text-Fabric can be used to query the *BHSA* syntax data or to navigate the linguistic objects via Python loop functions. For example, a simple Python script can iterate through all phrase objects in the *BHSA*, stopping at each one to check whether its features match a set of conditions, and saving the object references which do. In Python, one can easily manipulate strings, count objects, and apply advanced statistical conversions to the counts. The *BHSA* data is preprocessed for this project in two ways. First, cardinal number chains are segmented (or "chunked") so that long quantifier noun phrases can be easily isolated. In the same way, stacked prepositions (e.g. מִקֵּץ) are chunked. Second, in some cases the *BHSA* separates time phrases into two separate phrases, even though they are functionally one. Those cases are merged into new objects which are used for the analysis.[[74]](#footnote-74)

The approach of this analysis is data-driven and inductive, meaning that few semantic categories are assumed. This method is likewise connected with usage-based language theory. The model predicts that forms which are very common also function as linguistic prototypes through which rare or novel uses are construed.[[75]](#footnote-75) Indeed, raw counts of language units often produce only a few forms which are used very frequently. When their frequency is plotted against their rank, a sharp downward curve can be seen, which is associated with a principle known as Zipf's law.[[76]](#footnote-76) It is hypothesized that the same curve in the time data represents the prototypical forms used for construing Hebrew time adverbials. As such, it will be shown that rarer forms often exploit common structures by extending or modifying them.[[77]](#footnote-77) Another principle which guides the analysis is that of semantic association, which states that forms which are more associated semantically will co-occur more frequently.[[78]](#footnote-78) The goal herein is to use the prototype and collocation data to construct data-driven semantic classes for time adverbials. These classes can then be used in further analysis for word order and sentence argumentation.

A final word is necessary about the statistical methodology. It is common in Hebrew studies to use raw counts of a form. However, raw counts can hide surprising differences in distribution. For example, if a given form A occurs 9 of 10 times with another form X (90%), it is more significant than B's 18 of 100 times with X (18%). Corpus linguists have recognized that corpus size should also play a role in determining significance.[[79]](#footnote-79) In total, four different counts are relevant in determining how significant a given co-occurrence count is: [[80]](#footnote-80)

# Figure : Contingency Table for Co-occurrence Data

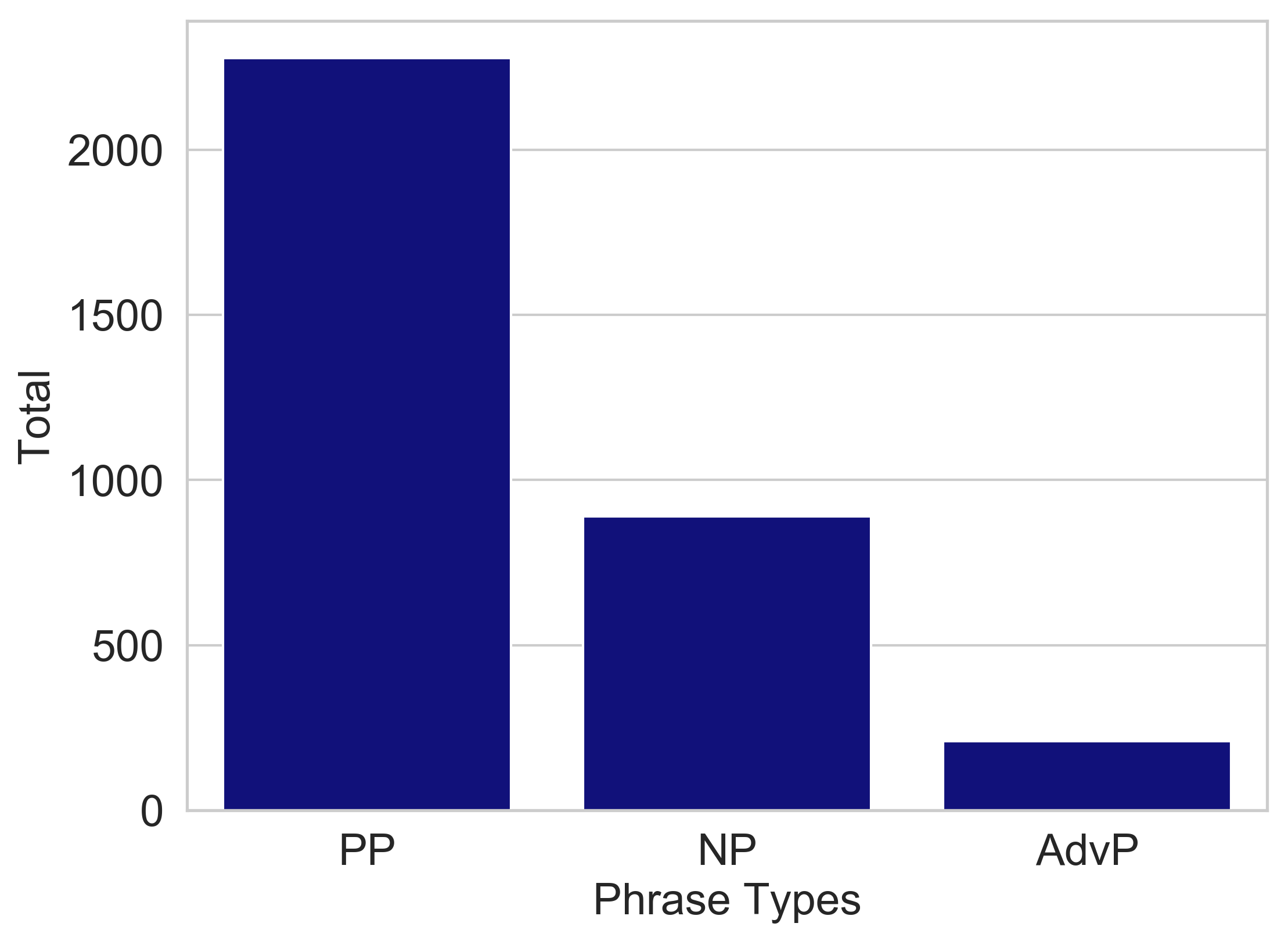
|  |  |  |
| --- | --- | --- |
|  | **target form** | *all other target forms* |
| **co-occurring form X** | A | B[[81]](#footnote-81) |
| *all other co-occurring forms* | C | D |

Given these four counts, a number of tests can be applied which measure how statistically significant a co-occurrence is. This project uses log-transformed Fisher's Exact scores, which are ideal for non-normal distributions and low values.[[82]](#footnote-82)

The *BHSA* contains a total of 249,383 Hebrew phrases, with 3,961 of them possessing an adverbial time label. The total comes to 3,376 after merging superfluous time phrases and removing cases headed by אָז, עַתָּה, כֵּן, and אַךְ. These phrases constitute the starting point for the analysis. The distribution of formal types is shown in the table and plot below.

# Figure : Time Phrase Types

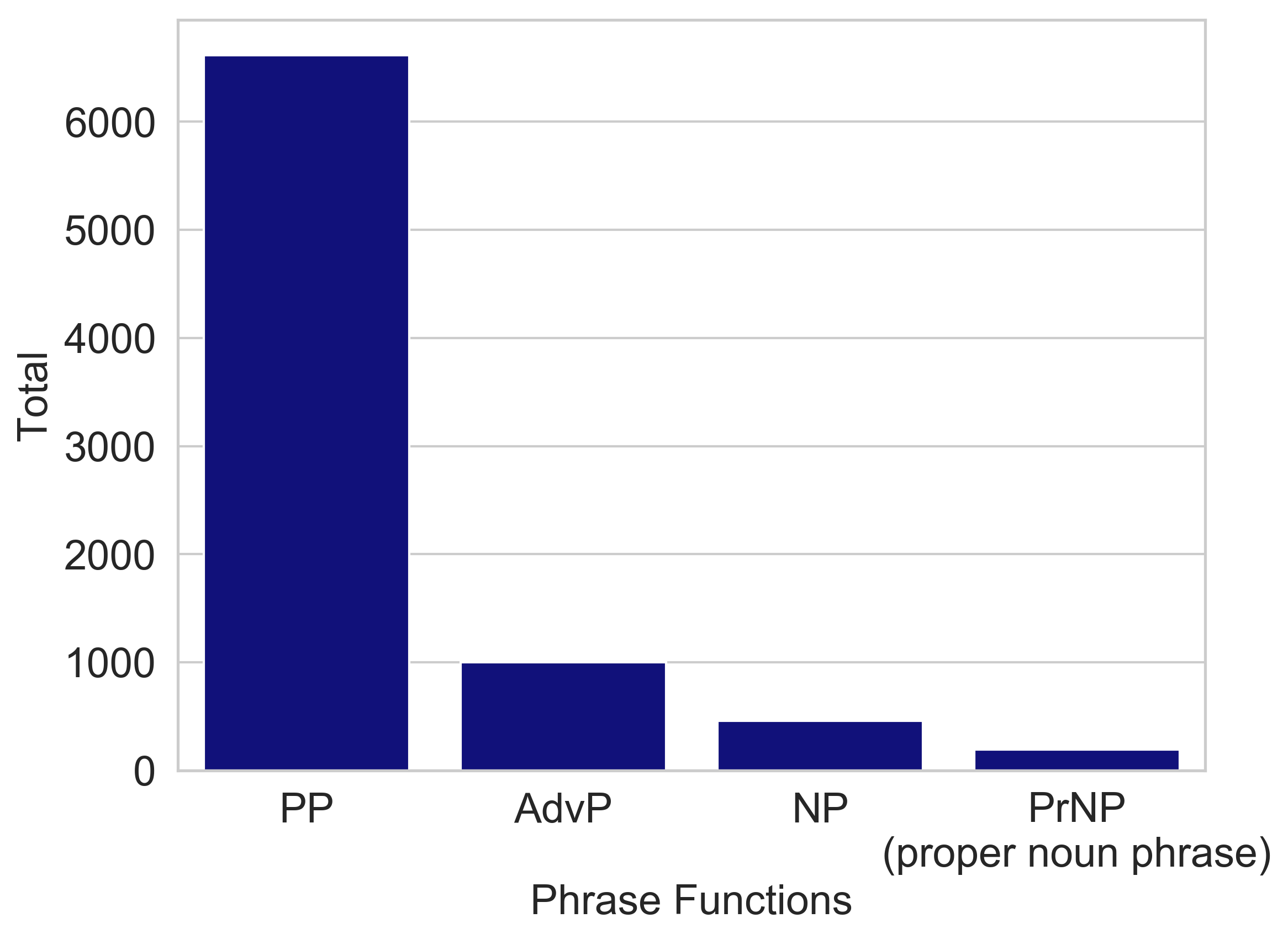
|  |  |
| --- | --- |
|  | Total |
| PP | 2277 |
| NP | 890 |
| AdvP | 209 |



Note that the adverb phrase (advp) has 421 less cases due to the removal of the particles. Ultimately, there are syntactic differences between the np and advp, as will be shown. The prepositional phrase (pp) dominates time with 2,277 or 67% of all forms. The distribution is comparable to, but lower than, location, with 80% pp:

# Figure : Location Phrase Types[[83]](#footnote-83)

|  |  |
| --- | --- |
|  | Total |
| PP | 6612 |
| AdvP | 1002 |
| NP | 461 |
| PrNP | 191 |



The difference in pp distribution between location and time is statistically significant, with an association score of -45, showing negative association, where significance is any score > 1.3 or < -1.3 .[[84]](#footnote-84) The attraction of np to time is even greater, with a score of 198 (showing strongly positive association). Thus, the pp appears to play a slightly smaller role in time. The np, on the other hand, plays a larger role in time semantics. The cause of this difference is the special use of the zero-marked time np for durative function, to be discussed.[[85]](#footnote-85) Despite the differences, predominance of pp in both time and location offers a clue that Biblical Hebrew, like other languages, encodes the two similarly.[[86]](#footnote-86)

The table below shows the top 50 semantic heads throughout all time adverbials. The definition of head adopted herein is that of Croft's "primary information bearing unit," i.e. "content words."[[87]](#footnote-87)

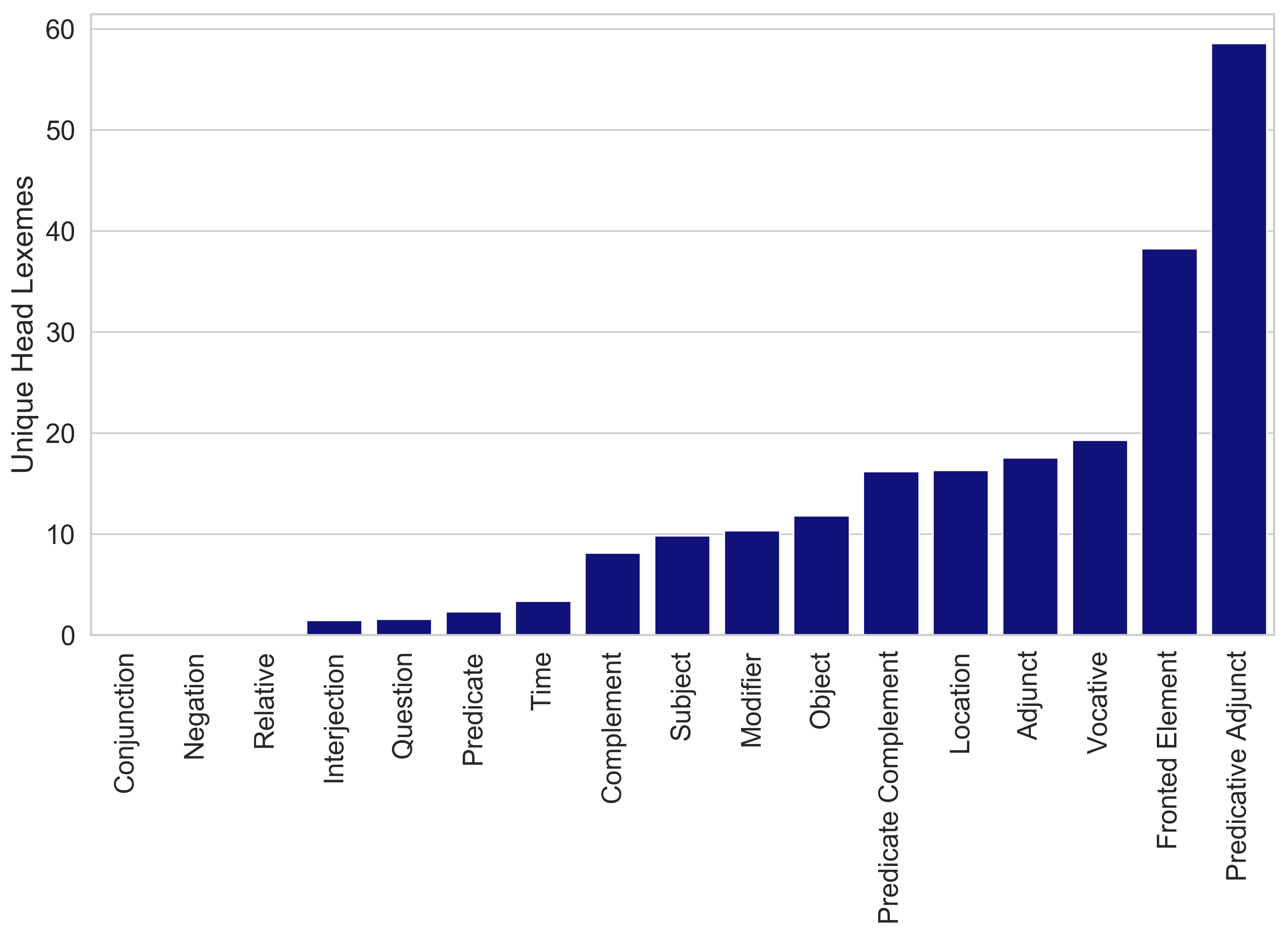
# Figure : Top 50 Semantic Heads for Time Adverbials

|  |  |
| --- | --- |
|  | Total |
| יֹום | 1524 |
| שָׁנָה | 376 |
| עֹולָם | 197 |
| עֵת | 165 |
| לַיְלָה | 139 |
| בֹּקֶר | 133 |
| עֶרֶב | 82 |
| חֹדֶשׁ | 79 |
| אַחַר | 67 |
| יֹומָם | 38 |
| מָחָר | 38 |
| תָּמִיד | 33 |
| מָוֶת | 31 |
| נֵצַח | 28 |
| מָתַי | 25 |
| מָחֳרָת | 25 |
| פָּנֶה | 23 |
| דֹּור | 21 |
| עַד | 20 |
| רִאשֹׁון | 18 |
| תְּמֹול | 17 |
| צָהֳרַיִם | 14 |
| מֹועֵד | 14 |
| דָּבָר | 14 |
| אָן | 14 |
| אֶחָד | 11 |
| שַׁבָּת | 9 |
| רֶגַע | 8 |
| נְעוּרִים | 8 |
| חַיִּים | 7 |
| שָׁבוּעַ | 7 |
| אֶתְמֹול | 6 |
| הֵנָּה | 6 |
| בֶּטֶן | 5 |
| שְׁבִיעִי | 5 |
| נֶשֶׁף | 5 |
| רֹאשׁ | 5 |
| אַרְבַּע | 5 |
| רֵאשִׁית | 5 |
| תְּחִלָּה | 5 |
| יֶרַח | 5 |
| אֶמֶשׁ | 4 |
| שָׁלֹשׁ | 4 |
| מַלְכוּת | 4 |
| תְּשׁוּבָה | 4 |
| זֹאת | 4 |
| קֶדֶם | 4 |
| אַחֲרֹון | 4 |
| קַיִץ | 4 |
| טֶרֶם | 4 |

As recognized by DeVries and Brin, יוֹם is by far the most common head in Biblical Hebrew time adverbials.[[88]](#footnote-88) In addition to יוֹם, other natural/calendrical times appear such as שָׁנָה, לַיְלָה, בֹּקֶר, עֶרֶב, חֹדֶשׁ, צָהֳרַיִם, נֶשֶׁף, יֶרַח, and קַיִץ. These units constitute what Haspelmath calls the "canonical" times, which constitute the most prevalent units across world languages.[[89]](#footnote-89) These times are reference points known from "the human natural environment on earth," deriving from "the alternation of light and dark, changes in the shape of the moon, and changes in the path of the sun across the sky."[[90]](#footnote-90) Other terms are more culturally defined, such as שַׁבָּת and שָׁבוּעַ. The presence of various nouns not intuitively associated with time include מָוֶת, נְעוּרִים, חַיִּים, בֶּטֶן, and מַלְכוּת. These suggest reference points related to the life cycles or reigns of individuals. Cardinal numbers such as אֶחָד, אַרְבַּע, and שָׁלֹשׁ, as well the ordinal שְׁבִיעִי, suggest the use of positions along a number line as reference points.[[91]](#footnote-91) The list also contains more abstract nouns such as עוֹלָם, עֵת, מָחָר, פָּנֶה (positional use), נֵצַח, דּוֹר, תְּמוֹל, מוֹעֵד, רֶגַע, אֶתְמוֹל, רֵאשִׁית, תְּחִלָּה, אֶמֶשׁ, תּשׁוּבָה, קֶדֶם, אַחַרוֹן, and טֶרֶם, terms that designate various time points, durations, or inherent positions. Two interrogative markers are also present, מָתַי and אָן, as well as the deictic particle הֵנָּה.

The intuition that the heads of time are semantically specialized can be operationalized and tested by comparing how diverse time heads are versus other arguments. The chart below demonstrates that time is more comparable to the predicate than location or other adjuncts.

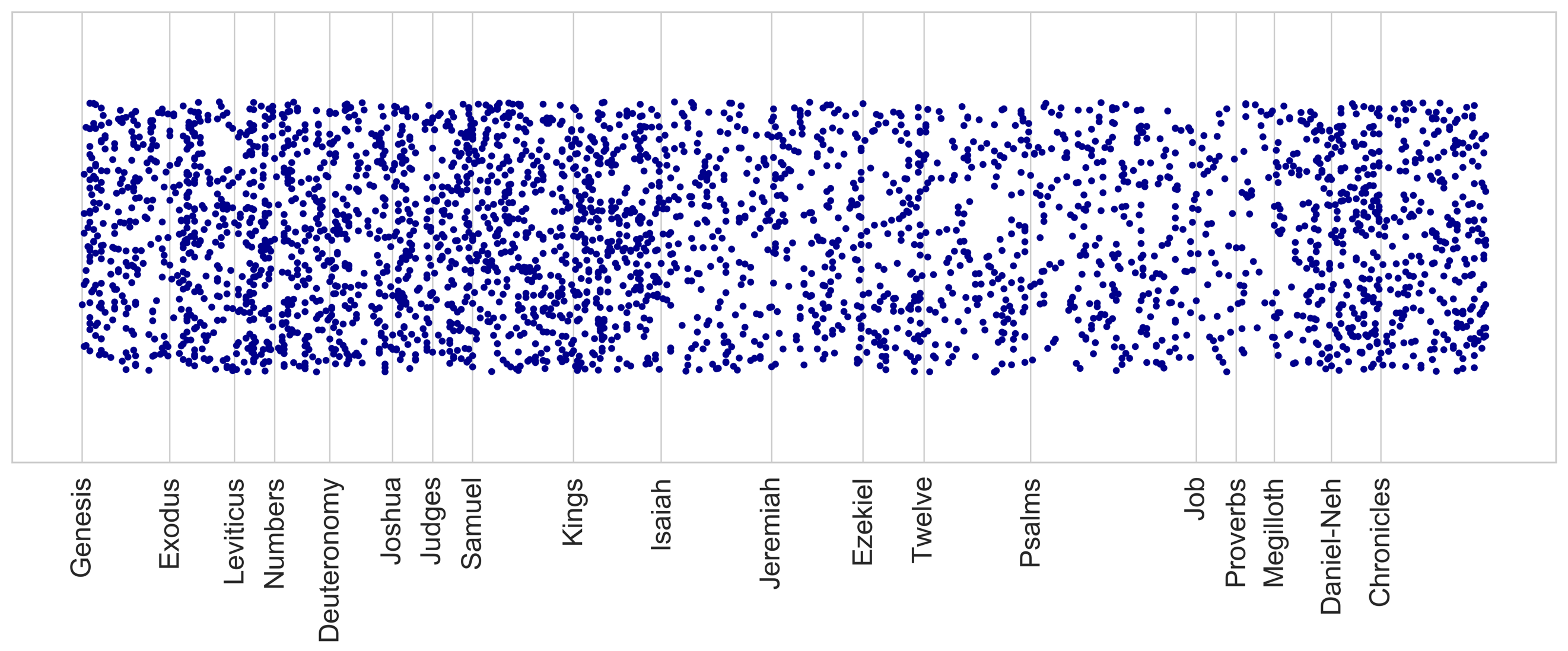
# Figure : Unique Head Lexemes for Sentence Arguments Per 100 Uses[[92]](#footnote-92)



The data thus gives empirical backing for the sense that time words are lexicalized specifically for time function.[[93]](#footnote-93) It is also intriguing that time is as selective as predicates, given the close relation between temporality and verbal action.

The idea that time adverbials are "optional" might give the impression that they are less than vital language components.[[94]](#footnote-94) The distribution of time adverbials paints a different picture. This strip chart visualizes time distribution from chapter to chapter throughout the Hebrew Bible.

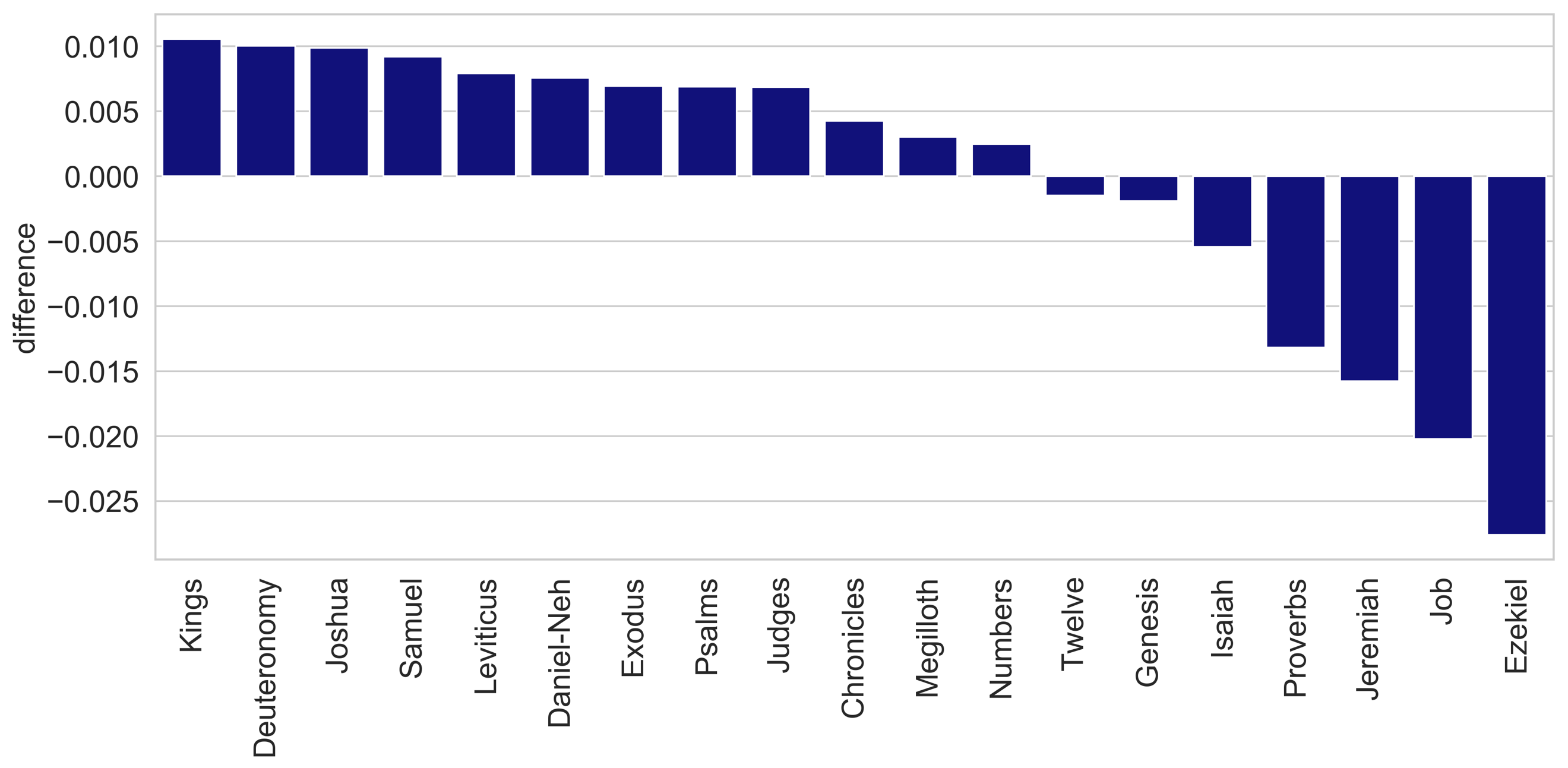
# Figure : Adverbial Time Distribution from Chapter to Chapter throughout the Hebrew Bible (Y-axis is random)



time adverbials are distributed throughout the whole corpus, though there are differences in concentration. Notably, time enjoys the strongest attestation in narrative texts. This confirms the intuition amongst Hebraists that time adverbials serve a special role in anchoring and directing narratives.[[95]](#footnote-95)

While the strip chart offers a heuristic picture, a more precise measure of distribution is needed to directly compare distributions. The concept of Deviation of Proportion can tell whether a given book is overrepresented or underrepresented in its frequency of time adverbials. The value is based on a corpus part's proportional representation.[[96]](#footnote-96) For example, Genesis accounts for 7% of all phrases in the Hebrew Bible. Thus, based on a random distribution, it is expected to account for 7% of all time adverbials. Anything more or less shows deviation from the expected proportion. Applying this formula allows for a more precise look at which time distribution.

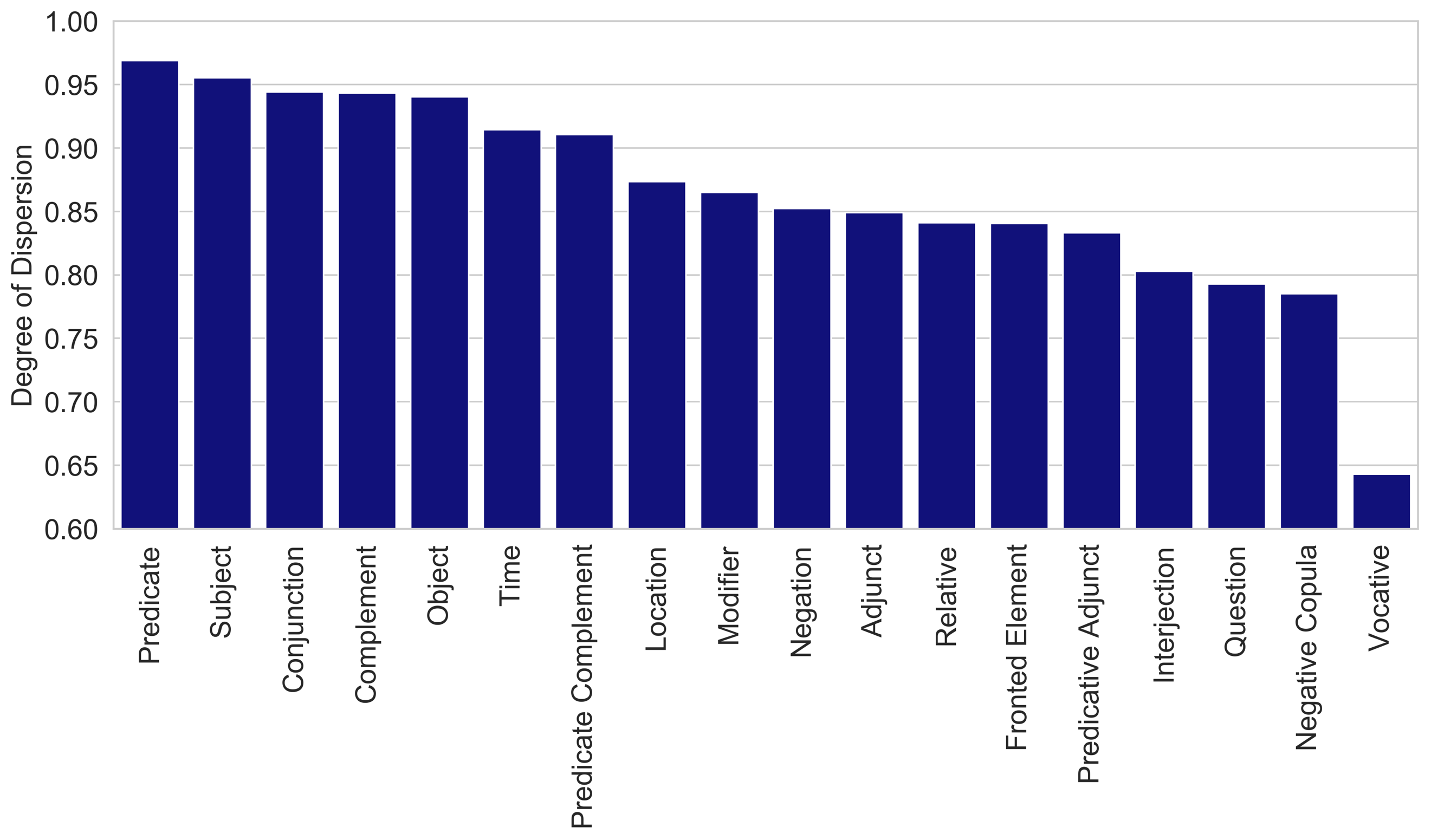
# Figure : Deviation of Proportions from a Random Distribution for Time Adverbials per Book (grouped)



Consistent with the strip chart, the narrative books show a slightly higher preference for time than do poetic books. Kings has the greatest, with 1% higher than expected, while Ezekiel, the least, has 3% (0.0275) *less* than expected. Exploring why these disparities exist, and whether they have any functional implications for time adverbials is an area to be explored in the full project.

It is also possible to compare the distribution of time against other sentence arguments. Degree of Dispersion is a statistical measure for the distribution of forms throughout the various parts of a corpus.[[97]](#footnote-97) It uses the Deviation of Proportion measure as part of a formula to quantify evenness of distribution across parts of a corpus. The result is a decimal value, from 0-1, that describes how well distributed a form is relative to a random spread, with 1 being perfectly even and 0 being completely uneven (theoretically). [[98]](#footnote-98) The chart below shows how the distribution of time compares with other major sentence arguments encoded in *BHSA*.

# Figure : Degree of Dispersion for Sentence Arguments in BHSA (higher is more evenly distributed)



The data shows that time (0.91) is comparable in distribution to key arguments such as object (0.94) and is slightly more evenly distributed than location markers (0.87). The difference between time and other kinds of adjuncts (0.85) is even greater, with a difference of 0.07 (7%). This data, alongside the data showing time's preference for specialized heads, reveals the time adverbial to be an important, dedicated component in Biblical Hebrew. "Optional" it may be from the perspective of verb meaning,[[99]](#footnote-99) but it cannot be so for the language as a whole.

Many time heads are governed by a further preposition. 67% of all time adverbials prepositional phrases (2,277). The prepositions are often "stacked." The chart below shows all of the attested prepositional combinations and their respective raw counts. Note that prepositions have been lemmatized and stripped of accents to disambiguate minor differences. Stacked prepositions are dot-separated.

# Figure : Attested Prepositions and Preposition Chains in Time Adverbial pps

|  | Total |
| --- | --- |
| ב | 1269 |
| עד | 314 |
| ל | 294 |
| מן | 155 |
| אחר | 84 |
| כ | 49 |
| מן.קץ | 19 |
| ב.אחרית | 15 |
| ל.מן | 14 |
| ב.עוד | 11 |
| מן.קצה | 10 |
| ל.פנה | 9 |
| בין | 8 |
| ל.קץ | 5 |
| עד.ל | 3 |
| את | 2 |
| על | 2 |
| ב.תוך | 2 |
| ל.פנה.מן | 1 |
| ל.ראשׁ | 1 |
| עד.אחר | 1 |
| מן.קצת | 1 |
| עד.מן | 1 |
| עד.בלת | 1 |
| מן.ל | 1 |
| אל | 1 |
| ב.ראשׁ | 1 |
| כ.מן | 1 |
| בלת | 1 |
| ל.מן.ב | 1 |

בְּ dominates the list, by itself accounting for 56% of all uses. The list contains the expected "canonical" prepositions עַל, לְ, מִן, כְּ, בַּיִן. The count also includes functional prepositions such as אַחַר, קֵּץ, אַחֲרִית, פָּנֶה, תָּוֶךְ, רֹאשׁ, and קְצָת. The list includes two modifiers which are not strictly prepositions but rather adverbial modifiers, עוֹד and בֵּלֶת; representing a small number of cases (13) these are presently coded this way to aid in semantic head selection.[[100]](#footnote-100) There are a number of surprises: אֵת, עַל, and אֶל. These are reviewed later.

[Role of Constituents]

[Primitive clusters]

[Generalizations — the role of associations]

[Order experiment]

[verb collocation experiment]

1. E.g. Ken Penner, “Verbal System, History of the Research,” in *Encyclopedia of Hebrew Language and Linguistics*, ed. Geoffrey Khan, vol. 3 (Leiden: Brill, 2013), 918–921; John A. Cook, “Current Issues in the Study of the Biblical Hebrew Verbal System,” *Kleine Untersuchungen zur Sprache des Alten Testaments und seiner Umwelt* 17 (2014): 79–108; Leslie McFall, *The Enigma of the Hebrew Verbal System: Solutions from Ewald to the Present Day*, Historic Texts and Interpreters in Biblical Scholarship (The Almond Press, 1982). [↑](#footnote-ref-1)
2. "This is so much so that it creates the impression that BH did not have any conventions in this regard that were unique to it and that could be used to express significant [temporal] nuances in a BH narrative." Christo H.J. Van der Merwe, “Reconsidering Biblical Hebrew Temporal Expressions,” *ZAH* 10, no. 1 (1997): 42. [↑](#footnote-ref-2)
3. E.g. Wolfgang Klein, *Time in Language*, Germanic Linguistics (London: Routledge, 1994); Martin Haspelmath, *From Space to Time: Temporal Adverbials in the World’s Languages*, LINCOM Studies in Theoretical Linguistics 2 (Münchn: Lincom Europa, 1997). [↑](#footnote-ref-3)
4. Christo H.J. Van der Merwe, “‘Reference Time’ in Some Biblical Temporal Constructions,” *Biblica* 78, no. 4 (1997): 503–524; Geoffrey Khan, “Extraposition and Pronominal Agreement in Semitic Languages” (SOAS University of London, 1984), 79. [↑](#footnote-ref-4)
5. Jan Joosten, *The Verbal System of Biblical Hebrew*, Jerusalem Biblical Studies 10 (Jerusalem: Simor Ltd., 2012), 95–123; John A. Cook, *Time and the Biblical Hebrew Verb: The Expression of Tense, Aspect, and Modality in Biblical Hebrew*, Linguistic Studies in Ancient West Semitic 7 (Winona Lake: Eisenbrauns, 2012), 281–283. [↑](#footnote-ref-5)
6. C. H. J. Van der Merwe, J. A. Naudé, and Jan Kroeze, *A Biblical Hebrew Reference Grammar*, Second edition. (New York: Bloomsbury T&T Clark, 2017), 427, 429; Alexey Lyavdansky, “Temporal Deictic Adverbs as Discourse Markers in Hebrew, Aramaic and Akkadian,” *Journal of Language Relationship* 3 (2010): 22–42. [↑](#footnote-ref-6)
7. The closest are the works of De Vries and Brin, to be reviewed below. Simon J. DeVries, *Yesterday, Today and Tomorrow: Time and History in the Old Testament* (USA: Eerdmans, 1975); Gershon Brin, *The Concept of Time in the Bible and the Dead Sea Scrolls*, Studies on the Texts of the Desert of Judah 39 (Leiden: Brill, 2001). [↑](#footnote-ref-7)
8. See footnote 12. [↑](#footnote-ref-8)
9. E.g. Adele E. Goldberg, Devin M. Casenhiser, and Nitya Sethuraman, “Learning Argument Structure Generalizations,” *Cognitive Linguistics* 15, no. 3 (2004): 289–316; Susan Hunston and Gill Francis, *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*, vol. 4, Studies in Corpus Linguistics (Amsterdam: John Benjamins, 2000). [↑](#footnote-ref-9)
10. Adele E. Goldberg, *Constructions: A Construction Grammar Approach to Argument Structure* (Chicago: University of Chicago Press, 1995). [↑](#footnote-ref-10)
11. Anatol Stefanowitsch and Stefan Th. Gries, “Collostructions: Investigating the Interaction of Words and Constructions,” *International Journal of Corpus Linguistics* 8, no. 2 (2003): 209–243. [↑](#footnote-ref-11)
12. E.g. Natalia Levshina and Kris Heylen, “A Radically Data-Driven Construction Grammar: Experiments with Dutch Causative Constructions,” in *Extending the Scope of Construction Grammar*, ed. Ronny Boogaart, Timothy Colleman, and Gijsbert Rutten, Cognitive Linguistics Research 54 (Berlin: De Gruyter Mouton, 2014), 17–46. [↑](#footnote-ref-12)
13. Wilhelm Gesenius, Emiel Friedrich Kautsch, and A.E. Cowley, *Gesenius’ Hebrew Grammar*, Second. (Oxford: Clarendon Press, 1909), §100, 118; Paul Joüon and T. Muraoka, *A Grammar of Biblical Hebrew*, Subsidia Biblica 14/1-14/2 (Roma: Editrice Pontificio Istituto Biblio, 1996), §102a-c, 126i, 133; Bruce Waltke and M. O’Connor, *An Introduction to Biblical Hebrew Syntax* (Winona Lake: Eisenbrauns, 1990), §10.2.2c, 11.2, 39.3.1h; Carl Brockelmann, *Hebräische Syntax* (Neukirchen: Kreis Moers, Verlag der Buchhandliung des Erziehungsvereins, 1956), 92, 98–109; Bill T. Arnold and John H. Choi, *A Guide to Biblical Hebrew Syntax* (New York, N.Y: Cambridge University Press, 2003), 18–19, 103–110, 127; Ronald J. Williams and John C. Beckman, *Williams’ Hebrew Syntax*, 3rd ed. (Toronto: University of Toronto Press, 2007), 21, 137, 97–136; Jan Pieter Lettinga, *Grammaire de l’hébreu biblique* (Leiden ; Boston: Brill, 1999), 144, 177; Eduard König, *Historisch-Kritisches Lehrgebäude Der Hebräischen Sprache* (Leipzig: J.C. Hinrichs’sche Buchhandlung, 1895), 234, 262–265; Hans Bauer, Pontus Leander, and Paul Kahle, *Historische Grammatik Der Hebräischen Sprache Des Alten Testamentes* (Halle A.S.: M. Niemeyer, 1922), §80, 80l. [↑](#footnote-ref-13)
14. *Joüon*, §102b, 166k; *IBHS*, 39.3.1h. [↑](#footnote-ref-14)
15. See references in 12. Gibson, however, refuses the terminology of "accusative." John C. L. Gibson, *Davidson’s Introductory Hebrew Grammar Syntax*, 4. ed. (Edinburgh: Clark, 1994), 24. [↑](#footnote-ref-15)
16. *IBHS*, §8.1. [↑](#footnote-ref-16)
17. Van der Merwe, Naudé, and Kroeze, *BHRG*, §33.1; *Joüon*, 440. [↑](#footnote-ref-17)
18. *GKC*, §118; *Joüon*, §126; *IBHS*, §39.3.1d; Brockelmann, *Hebräische Syntax*, 92; Arnold and Choi, *A Guide*, 18–21; Gibson, *Davidson’s*, 140. [↑](#footnote-ref-18)
19. *GKC*, §118i; *Joüon*, §126i; *IBHS*, §10.2.2c. [↑](#footnote-ref-19)
20. *GKC*, §118i,k; *Joüon*, 126i; Van der Merwe, Naudé, and Kroeze, *BHRG*, §33.3; Brockelmann, *Hebräische Syntax*, 92; Arnold and Choi, *A Guide*, 19; Williams and Beckman, *Williams’*, 21. But see the more nuanced approach of Waltke and O'Connor, who divide into "deictics" and "independents." *IBHS*, 39.3.1h. [↑](#footnote-ref-20)
21. Van der Merwe, Naudé, and Kroeze, *BHRG*, §39; *GKC*, §119; *Joüon*, §133; *IBHS*, §11; Brockelmann, *Hebräische Syntax*, 98–109; Williams and Beckman, *Williams’*, 97–136. [↑](#footnote-ref-21)
22. Van der Merwe, Naudé, and Kroeze, *BHRG*, §39.1.4. [↑](#footnote-ref-22)
23. Ibid.; *GKC*, §119a; *IBHS*, §11.2d; Brockelmann, *Hebräische Syntax*, 92. [↑](#footnote-ref-23)
24. DeVries, *Yesterday*; Brin, *The Concept of Time*. [↑](#footnote-ref-24)
25. DeVries, *Yesterday*, 38. [↑](#footnote-ref-25)
26. Ibid. [↑](#footnote-ref-26)
27. E.g. see his concluding comments. Ibid., 335–350. [↑](#footnote-ref-27)
28. Brin, *The Concept of Time*, 25–51. [↑](#footnote-ref-28)
29. Ibid., 25. [↑](#footnote-ref-29)
30. Ibid., 52. [↑](#footnote-ref-30)
31. Ibid., 78–92. [↑](#footnote-ref-31)
32. Ibid., 58–64, 65–77. [↑](#footnote-ref-32)
33. Ibid., 95–124. [↑](#footnote-ref-33)
34. Ibid., 147–50. [↑](#footnote-ref-34)
35. Brin\* [↑](#footnote-ref-35)
36. Brin, *The Concept of Time*, 177–183. [↑](#footnote-ref-36)
37. Ibid., 184–185. [↑](#footnote-ref-37)
38. Ibid., 187. [↑](#footnote-ref-38)
39. Van der Merwe, “Reconsidering,” 1, 44–45. [↑](#footnote-ref-39)
40. Randolph Quirk, ed., *A Comprehensive Grammar of the English Language* (London ; New York: Longman, 1985), 528. [↑](#footnote-ref-40)
41. Janet Harkness, “Time Adverbials in English and Reference Time,” in *Essays on Tensing in English*, ed. Alfred Schopf, Linguistische Arbeiten 185, 228 (Tübingen: Max Niemeyer Verlag, 1987), 71–110. [↑](#footnote-ref-41)
42. Van der Merwe, “Reconsidering,” 48–49. [↑](#footnote-ref-42)
43. Ibid., 49. [↑](#footnote-ref-43)
44. Ibid., 45, 47. [↑](#footnote-ref-44)
45. E.g. *IBHS*, 49, 42–80. [↑](#footnote-ref-45)
46. E.g. Cynthia L. Miller-Naudé and Jacobus A. Naudé, “A Re-Examination of Grammatical Categorization in Biblical Hebrew,” in *From Ancient Manuscripts to Modern Dictionaries: Select Studies in Aramaic, Hebrew and Greek*, ed. Society of Biblical Literature, Tarsee Li, and Keith D. Dyer, Perspectives on linguistics and ancient languages 9 (Piscataway, NJ: Gorgias Press, 2017), 273–308. [↑](#footnote-ref-46)
47. \*substantiate [↑](#footnote-ref-47)
48. \*substantiate [↑](#footnote-ref-48)
49. William Croft, *Radical Construction Grammar: Syntactic Theory in Typological Perspective* (Oxford ; New York: Oxford University Press, 2001), 29–33. See also Martin Haspelmath, “Pre-Established Categories Don’t Exist: Consequences for Language Description and Typology,” *Linguistic Typology* 11 (2007): 119–132. [↑](#footnote-ref-49)
50. Goldberg, *Constructions*, 6–7; Croft, *Radical Construction Grammar*, 14–18. [↑](#footnote-ref-50)
51. For languages with productive adverb morphology, such as in English with -*ly*, adverbs are associated with the form. And just as novel words can be paired together with English -*ly* (e.g. "*That was done quite Cambridgely*"), Hebrew can use new words in adverb patterns. A specific example of this phenomenon is reviewed below with the "adverb" use of יוֹם. [↑](#footnote-ref-51)
52. "The constructional tail has come to wag the syntactic dog: everything from words to the most general syntactic and semantic rules can be represented as constructions." Croft, *Radical Construction Grammar*, 17. [↑](#footnote-ref-52)
53. Goldberg, *Constructions*, 7. [↑](#footnote-ref-53)
54. Derived from Goldberg with Hebrew examples added. See also Croft's version. Adele E. Goldberg, *Constructions at Work: The Nature of Generalization in Language*, Oxford linguistics (Oxford ; New York: Oxford University Press, 2006), 5; Croft, *Radical Construction Grammar*, 17. [↑](#footnote-ref-54)
55. *IBHS*, §39.3.1h. [↑](#footnote-ref-55)
56. Van der Merwe, Naudé, and Kroeze, *BHRG*, §39.14b. [↑](#footnote-ref-56)
57. Brin, *The Concept of Time*, 55. Obviously the extension to duration is motivated by the use of the plural. Though the derivation of constructions might be semantically motivated, their transformations are not always predictable. Goldberg, *Constructions at Work*, 69–72. [↑](#footnote-ref-57)
58. Van der Merwe, Naudé, and Kroeze, *BHRG*, §40.25. [↑](#footnote-ref-58)
59. "According to Construction Grammar, a distinct construction is defined to exist if one or more of its properties are not strictly predictable from knowledge of other constructions existing in the grammar." Goldberg, *Constructions*, 4. [↑](#footnote-ref-59)
60. Nick C. Ellis, Matthew Brook O’Donnell, and Ute Römer, “Usage-Based Language: Investigating the Latent Structures That Underpin Acquisition: Usage-Based Language,” *Language Learning* 63 (March 2013): 25–51. [↑](#footnote-ref-60)
61. Goldberg, Casenhiser, and Sethuraman, “Learning.” [↑](#footnote-ref-61)
62. R. Xiao, “Collocation,” in *The Cambridge Handbook of English Corpus Linguistics*, ed. D. Biber and R. Reppen (Cambridge: Cambridge University, 2015), 106–124. [↑](#footnote-ref-62)
63. See the analysis below. [↑](#footnote-ref-63)
64. Natalia Levshina, *How to Do Linguistics with R: Data Exploration and Statistical Analysis* (Amsterdam: John Benjamins, 2015), 223–239. [↑](#footnote-ref-64)
65. Anatol Stefanowitsch, “Empirical Cognitive Semantics: Some Thoughts,” in *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*, ed. Dylan Glynn and Kertin Fischer, Cognitive Linguistics Research 46 (Berlin: De Gruyter Mouton, 2010), 355–380; Stefan Fuhs, “The Aspectual Coercion of the English Durative Adverbial,” in *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*, ed. Dylan Glynn and Kerstin Fischer, Cognitive Linguistics Research 46 (Berlin: De Gruyter Mouton, 2010), 137–154; Levshina and Heylen, “Experiments with Dutch Causative Constructions”; Stefan Th. Gries and Anatol Stefanowitsch, “Extending Collostructional Analysis: A Corpus-Based Perspective on ‘Alternations,’” *International Journal of Corpus Linguistics* 9, no. 1 (2004): 97–125. [↑](#footnote-ref-65)
66. *BHSA* stands for *Biblia Hebraica Stuttgartensia Amstelodamensis*. Dirk Roorda et al., *ETCBC/Bhsa* (Zenodo, 2019), accessed April 23, 2019, https://zenodo.org/record/2554324. [↑](#footnote-ref-66)
67. The check consisted of tagging 1040\* representative surface forms for accuracy, as well as the correction of problematic cases. See the process in \*notebook. [↑](#footnote-ref-67)
68. Dirk Roorda, “The Hebrew Bible as Data: Laboratory - Sharing - Experiences,” in *Clarin in the Low Countries*, ed. J. Odijk and A. Van Hessen (London: Ubiquity Press, 2017). \*statistics citations. [↑](#footnote-ref-68)
69. Collostruction Analysis has been designed specifically for the non-normal distributions present in natural language. It relies on the nonparametric Fisher's Exact association test. It also has the express purpose of describing associations between lexical and grammatical units. Stefanowitsch and Gries, “Collostructions.” [↑](#footnote-ref-69)
70. Haspelmath, *From Space to Time*; Charles Fillmore, “Mini-Grammars of Some Time-When Expressions in English,” in *Complex Sentences in Grammar and Discourse: Essays in Honor of Sandra A. Thompson*, ed. Joan L. Bybee, Sandra A. Thompson, and Michael Noonan (Amsterdam: Benjamins, 2002); William Croft, *Verbs: Aspect and Causal Structure* (Oxford: Oxford University Press, 2012). [↑](#footnote-ref-70)
71. "Atom" objects are parts of whole objects which have gaps due to other intervening objects. Cody Kingham and Wido Van Peursen, “The ETCBC Database of the Hebrew Bible,” *Journal for Semitics* 27, no. 1 (2018): 3–6. [↑](#footnote-ref-71)
72. See features > function in the documentation. Dirk Roorda et al., “Biblia Hebraica Stuttgartensia (Amstelodamensis) Documentation,” Documentation, *ETCBC Github*, last modified April 10, 2019, accessed April 23, 2019, https://etcbc.github.io/bhsa/. [↑](#footnote-ref-72)
73. Dirk Roorda and Camil Staps, *Annotation/Text-Fabric: Sophisticated Data Retrieval from Repos* (Zenodo, 2019), accessed April 23, 2019, https://zenodo.org/record/2635046. [↑](#footnote-ref-73)
74. See the process at Cody Kingham, “Chunking (and Phrase Merges),” *Jupyter Notebook Viewer*, last modified April 23, 2019, accessed April 23, 2019, https://nbviewer.jupyter.org/github/CambridgeSemiticsLab/  
    BH\_time\_collocations/blob/master/analysis/preprocessing/chunking.ipynb. [↑](#footnote-ref-74)
75. Goldberg, Casenhiser, and Sethuraman, “Learning”; Ellis, O’Donnell, and Römer, “Usage-Based Language.” [↑](#footnote-ref-75)
76. Levshina, *Linguistics with R*, 62–68. [↑](#footnote-ref-76)
77. Adele Goldberg, *Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions* (Princeton, NJ: Princeton University Press, 2019). [↑](#footnote-ref-77)
78. Stefanowitsch and Gries, “Collostructions.” [↑](#footnote-ref-78)
79. Stefan Th. Gries, “Dispersions and Adjusted Frequencies in Corpora,” *International Journal of Corpus Linguistics* 13, no. 4 (2008): 403–437. [↑](#footnote-ref-79)
80. Adapted from Levshina, *Linguistics with R*, 223–224. [↑](#footnote-ref-80)
81. Note that B here is slightly different from the example provided, as it represents here *all* other targets besides A. [↑](#footnote-ref-81)
82. This is the method of Collostruction Analysis established by Stefanowitsch and Gries. Stefanowitsch and Gries, “Collostructions.” [↑](#footnote-ref-82)
83. These counts include locational complement arguments, which are isolated by selecting all complement phrases headed by a word that is statistically associated with the location argument. See that process in the following notebook. https://nbviewer.jupyter.org/github/CambridgeSemiticsLab/BH\_time\_collocations/blob/master/analysis/  
    preprocessing/head\_function\_associations.ipynb [↑](#footnote-ref-83)
84. The association score is the Fisher's Exact p-value with log10 transformation, as per Collostructional Analysis. A negative sign is added to scores that have a lower than expected frequency (compared to a random distribution). Stefanowitsch and Gries, “Collostructions.” [↑](#footnote-ref-84)
85. The zero-marked durative is the typical form cross linguistically. Haspelmath, *From Space to Time*, 120–126. [↑](#footnote-ref-85)
86. Ibid., 17–21. [↑](#footnote-ref-86)
87. Croft, *Radical Construction Grammar*, 258. See also Stephen L. Shead, *Radical Frame Semantics and Biblical Hebrew: Exploring Lexical Semantics*, BibInt 108 (Leiden: Brill, 2011), 104. [↑](#footnote-ref-87)
88. Note that the number of יוֹם reflected here, 1,524, is lower than Brin's count of 2,317 since this count only includes adverbial uses of time. Brin, *The Concept of Time*, 52. [↑](#footnote-ref-88)
89. Haspelmath, *From Space to Time*, 25–26. [↑](#footnote-ref-89)
90. Ibid., 25. [↑](#footnote-ref-90)
91. Note that these cases are not quantifier uses, since quantifier uses are excluded as semantic heads. [↑](#footnote-ref-91)
92. Give the statistic\*. [↑](#footnote-ref-92)
93. Van der Merwe, “Reconsidering,” 45. [↑](#footnote-ref-93)
94. E.g. Van der Merwe, Naudé, and Kroeze, *BHRG*, §33.1. [↑](#footnote-ref-94)
95. E.g. Van der Merwe, “Reference Time.” [↑](#footnote-ref-95)
96. Levshina, *Linguistics with R*, 82–85. [↑](#footnote-ref-96)
97. Gries, “Dispersions and Adjusted Frequencies,” 415–419. [↑](#footnote-ref-97)
98. Gries's original formula actually expresses the opposite, with 0 being perfectly distributed and 1 being uneven. However, I have adjusted the scores to more intuitively align evenness with size. This is done by simply subtracting each of Gries's original values from 1. [↑](#footnote-ref-98)
99. This assumption also needs to be tested. It is conceivable that for some verbs, especially cases like וַיְּהִי, the time adverbial is actually a necessary element. This is yet another area of investigation for the full project. [↑](#footnote-ref-99)
100. The final version of the project will optimize and reclassify these cases. [↑](#footnote-ref-100)