**Internal Collocations of Time Adverbials in Biblical Hebrew**

Cody Kingham

University of Cambridge

cak47@cam.ac.uk

**Abstract**

This article utilizes a cognitive-statistical method to analyze the composition of over 3,400 Biblical Hebrew (BH) phrases marked for adverbial time function. Adverbials in general, and time adverbials in particular, have received relatively little treatment in the literature. A cognitive-statistical method attempts to trace statistical relations in language back to cognitive relations, based on the hypothesis that cognitive categories produce constructional co-occurrences. This article applies this method to analyze the internal makeup of time adverbial phrases. This is done through two main experiments. The first details a statistical method to classify the parts of speech of time adverbial heads based on their collocational behavior. As expected, the investigation shows that time adverbial heads exhibit broadly binary behavior between 'adverb' and 'noun' part of speech categories. However, these tendencies are likewise marked by variability along a continuum, supporting a gradient approach to word classes. Second, this article tests which nominal modifier constructions are most strongly associated with time adverbial phrases. The associations are compared against 50,395 nominal phrases marked for time adjunct, location adjunct, (other) adjunct, complement, object, and subject roles in the ETCBC syntax database. The analysis finds that nominal time adverbials are stronger predictors for definite article modification than other arguments, due to the need for non-deictic anchoring of nominal time words. These findings shed light on the syntax and semantics of time adverbials while opening up new avenues for exploring linguistic categories with statistics.

**A Cognitive-Statistical Approach**

With the exception of a few small studies, phrasal time adverbials in Biblical Hebrew have received little systematic, linguistic treatment.[[1]](#footnote-1) The reason for this may be because simple phrase adverbials present no obvious challenges for the classical linguistic method. As the grammars reiterate, time adverbials are simply adverbs or noun phrases that indicate optional time reference.\* But this simple description leaves many unanswered questions. For instance, what qualifies as an adverb or a noun? How do time adverbial phrases differ or compare with other argument types? What do the individual components of time adverbials contribute to their semantic functions?[[2]](#footnote-2)

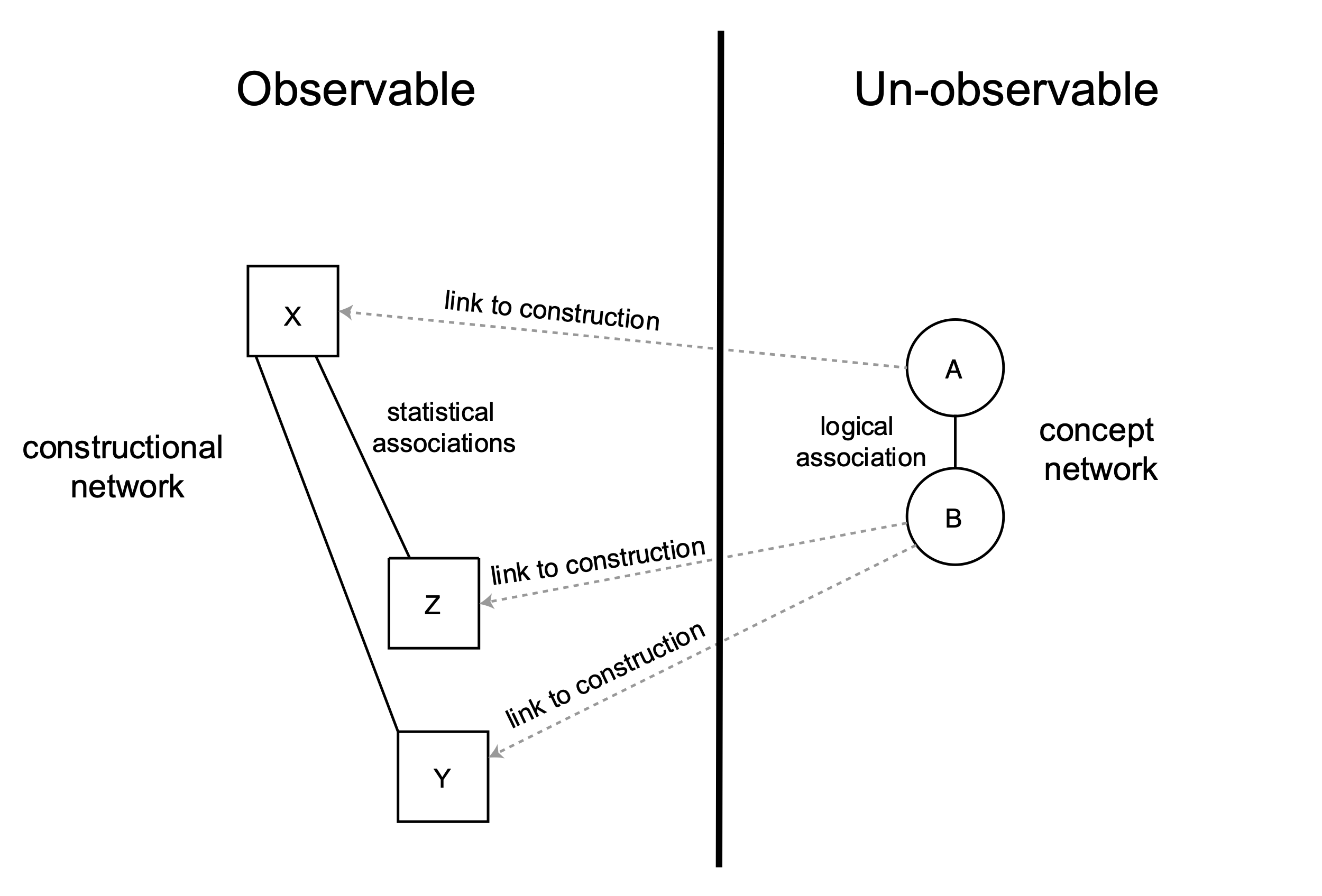
These questions are not trivial but go to the very foundation of grammatical inquiry. In recent years, Hebrew linguists have become more interested in foundational questions.[[3]](#footnote-3) The interest parallels a sea change in general linguistics, largely driven by new challenges from the cognitive linguistic perspective.[[4]](#footnote-4) At the same time, recent advancements in machine learning and brain imaging seem to indicate that language categories are learned rather than innate.[[5]](#footnote-5) This calls into question the degree to which *a priori* categories can be assumed within individual languages.[[6]](#footnote-6) In BH, linguists have long recognized that expected categories such as 'noun', 'adjective', and 'adverb' do not always fit.[[7]](#footnote-7) Given the new advancements, BH categories like 'adverb' and 'time adverbial' are ripe for fresh evaluation.

The cognitive approach provides an elegant and powerful explanation for the huge variety of forms observed across and within world languages. The simple human impetus to form categories around observed data drives the creation of linguistic patterns. Language, then, is a cognitive tool for categorizing the world.[[8]](#footnote-8) Humans connect categories built from experiential knowledge to linguistic sounds and signs (cf. Saussure).[[9]](#footnote-9) The concepts of prototypes ("best example"), family resemblance, and gradient categories are central to the structure of categories.[[10]](#footnote-10) As it turns out, these structures line up with quantitative patterns observed in the world. Long-tailed distributions (e.g. Zipf's Law), pervasive in natural, biological, and sociological systems, seem to drive the formation of prototypes.[[11]](#footnote-11) Hebraists will be familiar with this distribution from lexeme frequencies: a handful of terms (e.g. **וְ** 'and', **הַ** 'the', **אָמַר** 'say') occur very frequently, whereas thousands of terms occur only once (*hapax legomena*). Language learners likely use the most frequent items, the prototypes, as reference points for analogically learning new categories.[[12]](#footnote-12)

The link between statistical patterns and cognition opens the door to an empirical method for studying language categories.[[13]](#footnote-13) Mental associations between cognitive concepts give rise to statistical associations between language patterns.[[14]](#footnote-14) These statistical associations are seen most clearly in the phenomenon of cooccurrence, or collocation, between related constructions.[[15]](#footnote-15) For instance, **מֶלֶךְ** 'king' is statistically associated with **צִוָה**\* 'command' in the Hebrew Bible (HB).[[16]](#footnote-16) Conceptually this is due to the relatedness of the two terms with 'authority'. Synonym relations can be identified by indexing which terms collocate with the same kinds of items; for example: **שַׂר** 'prince'. This statistical concept has been successfully extended to grammatical constructions, further demonstrating that lexicon and syntax are not fundamentally distinct.[[17]](#footnote-17) The collocation principle underlies a number of recent advancements in machine translation.[[18]](#footnote-18)

The cognitive-statistical approach attempts to work backward from statistical associations towards hypothetical cognitive associations, as illustrated below.

Figure 1: Mutual relationship between linguistic, cognitive, and statistical data



While linguists have some introspective access to cognitive categories, it is unclear how reliable our own judgments can be.[[19]](#footnote-19) Linguists should therefore seek to operationalize their intuitions into testable hypotheses. They then act as observers standing on one side of a divide between empirical and experiential categories. This approach provides cognitive linguists with a more objective means for testing their hypotheses.[[20]](#footnote-20)

In BH, Forbes has pioneered data-driven models to study word class, using unsupervised clustering algorithms to automatically group words based on their co-occurrences.[[21]](#footnote-21) The approach yields promising results, with clear tendencies and recognizable groups. The methods of analysis for this study are different, but similarly reliant on the principle of collocation.

**Dataset**

The objective of this analysis is to measure the makeup and collocational tendencies of time adverbial phrases while comparing them with other argument types. To that end, the study requires a dataset that has wide coverage of the relevant syntactic categories and is freely accessible for scholarly scrutiny. Data-driven research is not theory neutral, nor is it immune to bias. This is equally true in the natural sciences. Rather, scientific research aims to make its assumptions clear, gather and measure data methodically, test those assumptions against the data, and make the results available for public scrutiny.[[22]](#footnote-22) Thus, any data-driven study that does not make its full dataset available does little to advance the field.

The open-sourced BHSA of the Eep Talstra Centre for Bible and Computer (ETCBC) meets the proposed criteria.[[23]](#footnote-23) The BHSA is accessed using Python and a corpus analysis package, Text-Fabric.[[24]](#footnote-24) It contains annotations for phrases interpreted to indicate adverbial time modification as well as numerous other argument roles, with coverage for the whole HB. The ETCBC built the annotations over a 40+ year history.[[25]](#footnote-25) As with any dataset, BHSA represents only one interpretation. It also possesses a few shortcomings. Importantly, relations below the phrase level are not always reliable, and there exists no tagging to explicitly indicate headship. For this reason, phrases must be pre-processed using a custom-built parser to isolate the necessary relations.[[26]](#footnote-26)

It is important to note that large scale corpus studies such as this should be measured against a different accuracy standard than traditional analyses. The retrieval and tagging of large quantities of text always assumes a tradeoff between precision and recall.[[27]](#footnote-27) Precision refers to the proportion of true positives in a query; recall refers to the proportion of true positives not missed by the query. Ideally the proportion would be 100% for both, but realistically large-scale studies require a compromise between perfection (technically impossible) and coverage.[[28]](#footnote-28) In this project, I have strived for a high degree of accuracy and coverage through manual inspection and iterative modifications. This was especially needed for the parsing of internal phrase relations. However, edge cases will always remain. Reliance on statistical significance helps to avoid erroneous conclusions. Also, the datasets have been published online as spreadsheets for scrutiny.[[29]](#footnote-29)

The primary dataset is a subset of all phrases in the Hebrew Bible stored in the BHSA. There are 99,426 phrases labeled as a subject, complement, object, location, or other adjunct (henceforth Subj, Cmpl, Objc, Loca, Adju). Many of these phrases contain rather complex relations, such as embedded sub-phrases, that are not always well marked in BHSA. To ensure accuracy, this set is pruned down to a set of shorter, simple phrases consisting of basic word relations that can easily be parsed. The resulting dataset contains 73,120 phrases that are then parsed for head words and modifiers using a generally semantic definition.[[30]](#footnote-30) A test to see whether the subsampled phrases negatively affected any particular book's representation showed no significant effects.[[31]](#footnote-31)

Table 1: Phrase function frequencies and ratio of selected samples versus BHSA[[32]](#footnote-32)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Adju | Cmpl[[33]](#footnote-33) | Loca | Objc | Subj | Time |
| *BHSA* | 9403 | 29568 | 2597 | 22394 | 31423 | 4041 |
| *sample* | 6737 | 17348 | 2125 | 16496 | 26156 | 3442 |
| *sample ratio* | 0.72 | 0.59 | 0.82 | 0.74 | 0.83 | 0.85 |

Throughout the study, I will refer to the phrase labels such as Time, Subj, etc. to distinguish the samples from the actual linguistic categories. Thus, I do not assume a one-to-one correspondence between the two. The tagged phrases can be viewed as tools of analysis that help to approach the object of study. This is on analogy with other scientific objects of measurement (e.g. a ruler, or a scale) which approximate rather than exhaust reality.[[34]](#footnote-34) Alternatively the phrases and tags can be viewed as a hypothesis (one interpretation) in need of testing.

**Parts of Speech Experiment**

As seen in the previous section, Time is headed by a variety of word classes. The grammars tend to treat time adverbials under three different rubrics: adverbs, adverbial accusatives, and temporal prepositional phrases.[[35]](#footnote-35) Adverbial accusatives and prepositional phrases are typically noun-based. Yet the distinction between an adverb and noun is not always clear, as this section will demonstrate. There appear to be both synchronic and diachronic reasons for the ambiguity. Synchronically, variant uses press against the neat and tidy categories once assumed by traditional grammars, suggesting that parts of speech exist on a gradient rather than on clear-cut bounds.[[36]](#footnote-36) Diachronically, adverbs can arise from a diverse set of sources, including nouns, adjectives, or specialized morphemes.[[37]](#footnote-37) BH itself lacks a productive morphology for adverbs, unlike languages such as English which possesses morphemes such as -ly.[[38]](#footnote-38) The most common source for adverbs in BH is nouns and adjectives.[[39]](#footnote-39) BH adverbs thus typically exist somewhere along an adverb-nominal continuum.

In order to study the variability found between BH time adverbial head words, a classification experiment is made to model gradient noun tendencies. The goal is simple: classify Time head lexemes based on their collocational behavior with nominalizing constructions. Nine types of noun modification are tagged in the sample: plural endings on the head, pronominal suffixes on the head, definite articles preceding the head,[[40]](#footnote-40) construct relations of the head (called 'genitive' here[[41]](#footnote-41)), appositional demonstratives or ordinal numbers, cardinal quantifiers, and qualitative quantifiers (e.g. **כֹל**).[[42]](#footnote-42) Another label, Ø, indicates no modifying specification of the headword except for the allowance of prepositions. The tagged features and their respective quantities for Time are shown below; features are allowed to co-occur.

Table 7: Collocational frequencies between Time head and nominal modifiers

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cardinal | definite | demonstrative | genitive | ordinal | plural | qualitative | suffix | Ø |
| 411 | 1331 | 493 | 426 | 162 | 569 | 212 | 103 | 1148 |

A separate table is compiled which tabulates how frequently a given head lexeme occurs with a given modifier. Words with a sample size less than 5 are dropped to ensure enough data is present for a reliable analysis. The resulting table contains 41 head lexemes. The counts for every head × modifier are then normalized to the total number of modifiers observed with a given head.[[43]](#footnote-43) The result is a decimal ratio from 0-1 (which can be read as a percentage when multiplied by 100), the proportion a given modifier represents of all a head's modifiers.

Table 8: Excerpt of Time head × modifier collocation frequencies

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| head | cardinal | definite | demonstrative | genitive | ordinal | plural | qualitative | suffix | Ø |
| יֹום | 165 | 819 | 383 | 279 | 108 | 384 | 171 | 41 | 9 |
| שָׁנָה | 188 | 42 | 7 | 22 | 25 | 82 | 4 | 2 | 0 |
| עַתָּה | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 354 |

Table 9: Excerpt of Time head × modifier collocation ratios

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| head | cardinal | definite | demonstrative | genitive | ordinal | plural | qualitative | suffix | Ø |
| יֹום | 0.07 | 0.35 | 0.16 | 0.12 | 0.05 | 0.16 | 0.07 | 0.02 | 0.0 |
| שָׁנָה | 0.51 | 0.11 | 0.02 | 0.06 | 0.07 | 0.22 | 0.01 | 0.01 | 0.0 |
| עַתָּה | 0.0 | 0.0 | 0.01 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.99 |

This multi-dimensional dataset would be tedious and error-prone to compare by hand (41×9 = 369 data points). Instead, an unsupervised clustering algorithm called Principle Component Analysis (PCA) is used to simultaneously compare all of the values and group them based on prevailing trends.[[44]](#footnote-44) The resulting graphs (Figure 6) should be interpreted with respect to space, where samples that are closer to one another are deemed more related based on the input features. Parts of speech categories from the lexicon are also added to the first graph to show how they correspond with the PCA classification.[[45]](#footnote-45) The second graph plots the same data but with the text of the heads, and with a slight adjustment of their positions for readability. The blue arrows in the plot show what has influenced each head's placement on the graph.[[46]](#footnote-46)

Figure 6: PCA analysis of head word tendencies

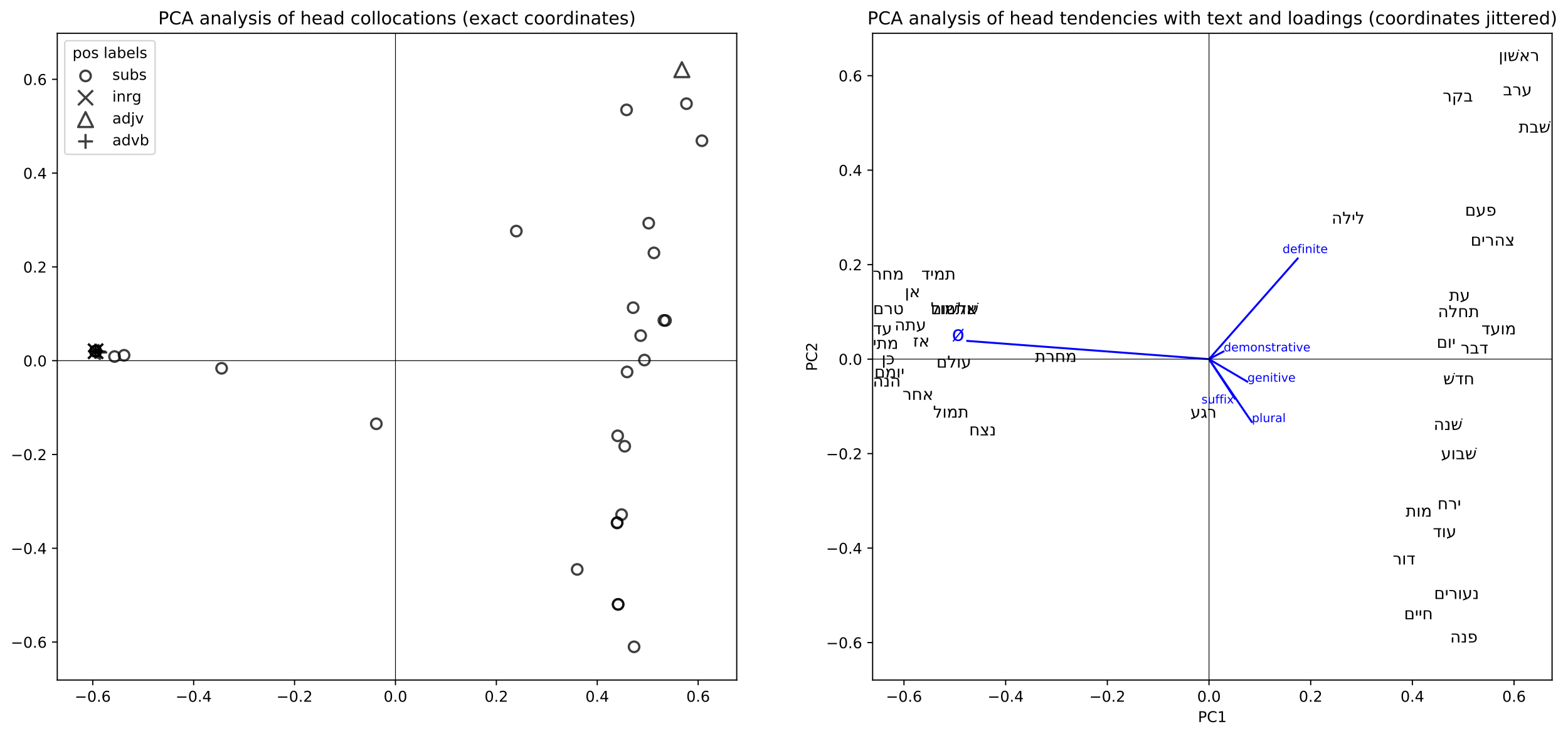
****

Table 10: Mean proportions for null (Ø) versus nominal modifiers by placement along x-axis

|  |  |  |
| --- | --- | --- |
|  | Ø | nominal |
| x<0 | 0.96 | 0.04 |
| x>0 | 0.04 | 0.96 |

The PCA graph reveals a clear separation of adverbs and nouns along a gradient. Table 5 shows that terms to the right of 0 occur with a nominal modifier on average 96% of the time, and on the left of 0 with null modification also with 96% on average. This again reflects the strong tendencies seen visually in the graph. In cognitive terms, the items on the extreme ends of the plot represent potential prototypes. The separation confirms the intuition that Time adverbial heads exhibit a broadly binary tendency between nouns and adverbs. Null (Ø) modification is the strongest contributor to the separation as seen by the long blue arrow.[[47]](#footnote-47)

The nominal side of the graph shows two subcategories separated by different modifier tendencies. The blue influences (loadings) such as definite + demonstrative, and genitive + plural + suffix demonstrate how PCA accurately identifies features which "cooperate" together in linguistic expressions. Definites and demonstratives frequently co-occur, for instance, in attributive adjunct expressions (e.g. **הַיוֹם הַזֶה**)[[48]](#footnote-48); plurals and suffixes co-occur in such constructions as **נְעוּרֵיךְ** 'your youth' or **חַיָּיו** 'his life'.[[49]](#footnote-49) It is noteworthy that these latter two terms, clustered along with **מוֹת** 'death' are thematically similar as life-events. This may explain their association with the suffix, which relates these terms to a personal entity.

While revealing clear prototypical behavior, the graph also shows fuzzy and intermediate tendencies. Firstly, it demonstrates the shortcoming of traditional part of speech values. Diverse lexicon tags of noun (subs), interrogative particles (inrg), and adverbs (advb) are found grouped together to the left of the y-axis. The right side is also mixed, containing one adjective (**רִאשׁוֹן** 'first').[[50]](#footnote-50) Secondly, several terms sit at intermediate positions along the x-axis, especially **רֶגַע** 'instant', **מָחֳרָת** 'next day', and **לַיְלָה** 'night'. Selected text examples for these terms will now be analyzed to understand their behavior.

The table below contains the full sample set for **רֶגַע**, a relatively small set (n=7). **רֶגַע** is placed slightly on the particle side of the graph.

Figure 7: Sampled sentences for *רֶגַע*

|  |  |  |  |
| --- | --- | --- | --- |
| reference | modifier | sentence | translation |
| *Num 17:10* | Ø | **וַאֲכַלֶּ֥ה אֹתָ֖ם כְּרָ֑גַע** | I will destroy them instantly. |
| *Isa 54:8* | Ø | **בְּשֶׁ֣צֶף קֶ֗צֶף הִסְתַּ֨רְתִּי פָנַ֥י רֶ֨גַע֙ מִמֵּ֔ךְ** | In a flood of anger I hid my face from you at an instant. |
| *Ps 73:19* | Ø | **אֵ֤יךְ הָי֣וּ לְשַׁמָּ֣ה כְרָ֑גַע** | How they are ruined instantly! |
| *Exod 33:5* | cardinal | **רֶ֧גַע אֶחָ֛ד אֶֽעֱלֶ֥ה בְקִרְבְּךָ֖** | At a single instant I will come up in your midst. |
| *Isa 26:20* | qualitative | **חֲבִ֥י כִמְעַט־רֶ֖גַע** | Hide instantly! [lit. "in a short instant"] |
| *Ezra 9:8* | qualitative | **כִּמְעַט־רֶגַע֩ הָיְתָ֨ה תְחִנָּ֜ה מֵאֵ֣ת׀ יְהוָ֣ה אֱלֹהֵ֗ינוּ** | In just a short instant our supplication is before YHWH our God. |
| *Isa 27:3* | plural | **לִרְגָעִ֖ים אַשְׁקֶ֑נָּה** | Continuously I watered it. |

3 of the text examples are used without any kind of nominal modifier, 3 employ quantifiers, and 1 has a plural ending. Given the small sample size, we should withhold strong judgments about this term, however a number of observations are pertinent. Even though this word has a slight majority of nominal modifiers (4/7), the algorithm has situated it on the adverb side due to the stronger influence of the null feature. Another reason is due to the weaker influence of quantifiers. This is not just an accidental artifact of the PCA analysis; the feature weights are determined across the whole dataset by the degree, and thus predictive power, of their variability. Furthermore, an abstract word like 'instant' does not seem to fit the semantics of a prototypical noun.[[51]](#footnote-51) The data thus tentatively indicate a term that sits somewhere on the border between a noun and adverb.

**מֳחָרָת** 'next day' is witnessed 26 times in the sample. There are 20 cases of null modification and 6 cases of nominal modification. Note that 11 null cases not shown are identical to the first example with **וַיְהִי**.[[52]](#footnote-52)

Table 11: Sampled modifiers with *מָחֳרָת*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cardinal | definite | demonstrative | genitive | ordinal | plural | qualitative | suffix | Ø |
| 0 | 1 | 0 | 5 | 0 | 0 | 0 | 0 | 20 |

Table 12: Selected samples of *מָחֳרָת*

|  |  |  |  |
| --- | --- | --- | --- |
| reference | modifier | sentence | translation |
| *Gen 19:34* | Ø | **וַֽיְהִי֙ מִֽמָּחֳרָ֔ת** | It was the subsequent day. |
| *1 Sam 5:4* | Ø | **וַיַּשְׁכִּ֣מוּ בַבֹּקֶר֮ מִֽמָּחֳרָת֒** | He arose early on the next morning. |
| *1 Sam 5:3* | Ø | **וַיַּשְׁכִּ֤מוּ אַשְׁדֹּודִים֙ מִֽמָּחֳרָ֔ת** | The Ashdodites arose early the next day. |
| *Judg 6:38* | Ø | **וַיַּשְׁכֵּם֙ מִֽמָּחֳרָ֔ת** | He arose the next day. |
| *Num 17:6* | Ø | **וַיִּלֹּ֜נוּ כָּל־עֲדַ֤ת בְּנֵֽי־יִשְׂרָאֵל֙ מִֽמָּחֳרָ֔ת עַל־מֹשֶׁ֥ה וְעַֽל־אַהֲרֹ֖ן** | On the next day the whole assembly of Israel grumbled against Moses and Aaron. |
| *Lev 19:6* | Ø | **בְּיֹ֧ום זִבְחֲכֶ֛ם יֵאָכֵ֖ל וּמִֽמָּחֳרָ֑ת** | On the day of your sacrifice it shall be eaten, and on the next day. |
| *Exod 32:6* | Ø | **וַיַּשְׁכִּ֨ימוּ֙ מִֽמָּחֳרָ֔ת** | They arose on the next day. |
| *Exod 9:6* | Ø | **וַיַּ֨עַשׂ יְהוָ֜ה אֶת־הַדָּבָ֤ר הַזֶּה֙ מִֽמָּחֳרָ֔ת** | YHWH did this very thing on the next day. |
| *Josh 5:12* | Ø | **וַיִּשְׁבֹּ֨ת הַמָּ֜ן מִֽמָּחֳרָ֗ת בְּאָכְלָם֙ מֵעֲב֣וּר הָאָ֔רֶץ** | The mannah ceased when they ate from the yield of the land. |
| *Josh 5:11* | genitive | **וַיֹּ֨אכְל֜וּ מֵעֲב֥וּר הָאָ֛רֶץ מִמָּֽחֳרַ֥ת הַפֶּ֖סַח מַצֹּ֣ות וְקָל֑וּי בְּעֶ֖צֶם הַיֹּ֥ום הַזֶּֽה׃** | They ate unleavened bread and roasted meat from the yield of the land on the day after the Passover, on that very day. |
| *Lev 23:16* | genitive | **עַ֣ד מִֽמָּחֳרַ֤ת הַשַּׁבָּת֙ הַשְּׁבִיעִ֔ת תִּסְפְּר֖וּ חֲמִשִּׁ֣ים יֹ֑ום** | Up to the day after the seventh Sabbath you shall count fifty days. |
| *Lev 23:11* | genitive | **מִֽמָּחֳרַת֙ הַשַּׁבָּ֔ת יְנִיפֶ֖נּוּ הַכֹּהֵֽן׃** | On the day after the Sabbath, the priest shall wave it. |
| *1 Sam 20:27* | genitive | **וַיְהִ֗י מִֽמָּחֳרַ֤ת הַחֹ֨דֶשׁ֙ הַשֵּׁנִ֔י** | It will be on the day after the second month. |
| *1 Chr 29:21* | genitive | **וַיַּעֲל֨וּ עֹלֹ֜ות לַיהוָ֗ה לְֽמָחֳרַת֮ הַיֹּ֣ום הַהוּא֒ פָּרִ֨ים אֶ֜לֶף אֵילִ֥ים אֶ֛לֶף כְּבָשִׂ֥ים אֶ֖לֶף וְנִסְכֵּיהֶ֑ם וּזְבָחִ֥ים לָרֹ֖ב לְכָל־יִשְׂרָאֵֽל׃** | On the day after that day they offered up as burnt offerings to YHWH a thousand bulls, a thousand rams, a thousand sheep, and their drink offerings, and their sacrifices in abundance for all of Israel. |
| *Jonah 4:7* | definite | **וַיְמַ֤ן הָֽאֱלֹהִים֙ תֹּולַ֔עַת בַּעֲלֹ֥ות הַשַּׁ֖חַר לַֽמָּחֳרָ֑ת** | God appointed a worm at the rising of the dawn of the next day. |

In 20 of 26 cases (77%), **מָחֳרָת** appears without nominal modification, reflecting adverb behavior similar to its lexical cousin **מָחָר** 'tomorrow' (Ø at 100%). There are 6 cases of explicit nominalization with 5 construct relations and 1 (vocalic) definite article. Specifically, the construct cases construe **מָחֳרָת** as a specific 'next day' after the Passover, the sabbath (2x), second month, and 'that day' (**הַיּוֹם הַהוּא**). The use with the (apparent[[53]](#footnote-53)) definite article, **לַמָּחֳרָת** 'at the next day' represents another nominalization of the term.

But perhaps unmodified cases of **מָחֳרָת** truly "are" just nouns used without modifiers? Yet, despite having a good sample size of 26, the behavior of **מָחֳרָת** differs significantly from the other regular nouns, resulting in its PCA placement. Furthermore, the explicitly nominalized cases of **מָחֳרָת** seem to interfere with an ongoing fusion to **מִן** 'since'. In nearly all cases, **מָחֳרָת** occurs with **מִן**, which appears to have been semantically bleached. The case in Lev 23:16 makes this clear with the double prepositions: **עַד מִמָּחֳרַת** 'until the day after'. Yet, in 2 of the 6 explicit nominal cases, **מָחֳרָת** appears without **מִן** (**הַיּוֹם הַהוּא** **לְמָחֳרַת** 'on the day after that day', 1 Chr 29:21; **לַמָחֳרַת** 'on the day after', Jonah 4:7). These 2 cases demonstrate that the noun construal may block an ongoing adverbialization. That is, as **מִמָּחֳרָת** grammaticalizes, it also becomes more semantically specialized for adverbial function, bringing it further away from a nominal sense. These cases without **מִן** may evidence such a transitional state.[[54]](#footnote-54)

לַיְלָה 'night' is the most frequent intermediate word in the graph, with 65% of its 131 sampled forms collocating with nominal modifiers. This is compared with an average nominal modification of 93% on the noun-side of the graph.

Table 13: Frequency and ratios of *לַיְלָה* with modification type

|  |  |  |
| --- | --- | --- |
|  | nominal | Ø |
| *frequency* | 85 | 46 |
| *ratio* | 0.65 | 0.35 |

Table 14: Sampled modifiers with *לַיְלָה* [[55]](#footnote-55)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cardinal | definite | demonstrative | genitive | ordinal | plural | qualitative | suffix | Ø |
| 2 | 81 | 19 | 0 | 0 | 5 | 19 | 0 | 46 |

Table 15: Selected samples of *לַיְלָה*

|  |  |  |  |
| --- | --- | --- | --- |
| reference | modifier(s) | sentence | translation |
| *Gen 40:5* | cardinal | **וַיַּֽחַלְמוּ֩ חֲלֹ֨ום שְׁנֵיהֶ֜ם אִ֤ישׁ חֲלֹמֹו֙ בְּלַ֣יְלָה אֶחָ֔ד** | The two of them each dreamed his own dream one night. |
| *Gen 41:11* | cardinal | **וַנַּֽחַלְמָ֥ה חֲלֹ֛ום בְּלַ֥יְלָה אֶחָ֖ד אֲנִ֣י וָה֑וּא** | He and I dreamed a dream one night. |
| *Gen 19:5* | definite | **אַיֵּ֧ה הָאֲנָשִׁ֛ים אֲשֶׁר־בָּ֥אוּ אֵלֶ֖יךָ הַלָּ֑יְלָה** | Where are the men who came in to you tonight? |
| *Exod 14:20* | qualitative+def. | **וְלֹא־קָרַ֥ב זֶ֛ה אֶל־זֶ֖ה כָּל־הַלָּֽיְלָה׃** | Neither one approached the other for the whole night. |
| *Judg 16:3* | qualitative+def. | **וַיִּשְׁכַּ֣ב שִׁמְשֹׁון֮ עַד־חֲצִ֣י הַלַּיְלָה֒** | And Samson laid down until the middle of the night. |
| *Ps 121:6* | definite | **יֹומָ֗ם הַשֶּׁ֥מֶשׁ לֹֽא־יַכֶּ֗כָּה וְיָרֵ֥חַ בַּלָּֽיְלָה׃** | By day the sun will not strike you, nor the moon at night. |
| *2 Chr 7:12* | definite | **וַיֵּרָ֧א יְהוָ֛ה אֶל־שְׁלֹמֹ֖ה בַּלָּ֑יְלָה** | YHWH appeared to Solomon in the night. |
| *Gen 19:35* | definite+demon. | **וַתַּשְׁקֶ֜יןָ גַּ֣ם בַּלַּ֧יְלָה הַה֛וּא אֶת־אֲבִיהֶ֖ן יָ֑יִן** | So also that night they gave their father wine to drink. |
| *2 Chr 1:7* | definite+demon. | **בַּלַּ֣יְלָה הַה֔וּא נִרְאָ֥ה אֱלֹהִ֖ים לִשְׁלֹמֹ֑ה** | On that night God appeared to Solomon. |
| *Isa 21:8* | qual. + def. + pl. | **וְעַל־מִ֨שְׁמַרְתִּ֔י אָנֹכִ֥י נִצָּ֖ב כָּל־הַלֵּילֹֽות׃** | I stand at my guard throughout all the nights. |
| *Song 3:1* | definite + plural | **עַל־מִשְׁכָּבִי֙ בַּלֵּילֹ֔ות בִּקַּ֕שְׁתִּי אֵ֥ת שֶׁאָהֲבָ֖ה נַפְשִׁ֑י** | Upon my bed in the nights I seek he whom my soul loves. |
| *Gen 14:15* | Ø | וַיֵּחָלֵ֨ק עֲלֵיהֶ֧ם׀ לַ֛יְלָה ה֥וּא וַעֲבָדָ֖יו | He and his servants were split up at night. |
| *Num 14:14* | Ø | וּבְעַמֻּ֣ד עָנָ֗ן אַתָּ֨ה הֹלֵ֤ךְ לִפְנֵיהֶם֙ יֹומָ֔ם  וּבְעַמּ֥וּד אֵ֖שׁ לָֽיְלָה׃ | In a pillar of cloud you walk before them by day, and in a pillar of fire at night. |
| *Judg 9:34* | Ø | וַיָּ֧קָם אֲבִימֶ֛לֶךְ וְכָל־הָעָ֥ם אֲשֶׁר־עִמֹּ֖ו לָ֑יְלָה | Abimelek arose, along with all the people with him, at night. |
| *2 Kgs 7:12* | Ø | וַיָּ֨קָם הַמֶּ֜לֶךְ לַ֗יְלָה | The king arose at night. |
| *Hos 4:5* | Ø | וְכָשַׁ֧ל גַּם־נָבִ֛יא עִמְּךָ֖ לָ֑יְלָה | Even the prophet will stumble with you at night. |
| *Neh 2:12* | Ø | וָאָק֣וּם׀ לַ֗יְלָה אֲנִי֮ וַאֲנָשִׁ֣ים׀ מְעַט֮ עִמִּי֒ | I arose at night along with the few men who were with me. |
| *2 Chr 35:14* | Ø | כִּ֤י הַכֹּהֲנִים֙ בְּנֵ֣י אַהֲרֹ֔ן בְּהַֽעֲלֹ֛ות הָעֹולָ֥ה וְהַחֲלָבִ֖ים  עַד־לָ֑יְלָה | For the priests, sons of Aaron, were offering up the burnt offering and the fat offering until night. |

Again there is the temptation to consider the null modifications of **לַיְלָה** as simple unmodified uses of a noun which are inherently definite. Indeed, the PCA analysis does place **לַיְלָה** on the nominal side of the graph, though in an intermediate position. Yet other terms like **יוֹם** 'day', **בֹּקֶר** 'morning', or **עֶרֶב** 'evening' which are likewise well-represented in the dataset do not have near the amount of null modified cases.[[56]](#footnote-56) **לַיְְלָה** clearly has a distinctive profile.

As with **מָחֳרָת**, there may be a diachronic reason for the mixed behavior of לַיְלָה tied up with adverbial specialization. Namely, it may be the case that the ָה at the end of לַיְלָה originates from the *heh locale* found in locative adverbs (e.g. **אֲרְצָה** 'to the land', **מִצְרָיִמָה** 'to Egypt'\*).[[57]](#footnote-57) The locative ending would locate in time rather than space, via semantic extension.\* Similar to **מָחֳרָת**, then, this may be a case of ongoing specialization towards the adverbial function which results in the loss of a prototypical noun sense.

Aside from intermediate words in the PCA plot, even those with decisive tendencies can be used in non-prototypical ways. **עוֹלָם** 'forever' occurs with Ø modification in 95.5% of its samples (170/178), with 8 exceptions:

Table 16: Selected nominalized samples of *עוֹלָם*

|  |  |  |  |
| --- | --- | --- | --- |
| reference | modifier | sentence | translation |
| *1 Kgs 8:13* | plural | **בָּנֹ֥ה בָנִ֛יתִי בֵּ֥ית זְבֻ֖ל לָ֑ךְ מָכֹ֥ון לְשִׁבְתְּךָ֖ עֹולָמִֽים׃** | I have surely built a lofty house for you, a place for your dwelling for eternity. |
| *Isa 45:17* | plural+genitive | **וְלֹא־תִכָּלְמ֖וּ עַד־עֹ֥ולְמֵי עַֽד׃ פ** | You will not be ashamed till eternity on. |
| *Jer 28:8* | definite | **הַנְּבִיאִ֗ים אֲשֶׁ֨ר הָי֧וּ לְפָנַ֛י וּלְפָנֶ֖יךָ מִן־הָֽעֹולָ֑ם** | The prophets who were before me and you since eternity past... |
| *Joel 2:2* | definite | **כָּמֹ֗הוּ לֹ֤א נִֽהְיָה֙ מִן־הָ֣עֹולָ֔ם** | There has not been one like it since eternity past. |
| *Ps 61:5* | plural | **אָג֣וּרָה בְ֭אָהָלְךָ עֹולָמִ֑ים** | I will dwell in your tent for eternity. |
| *Ps 77:8* | plural | **הַֽ֭לְעֹולָמִים יִזְנַ֥ח׀ אֲדֹנָ֑י וְלֹֽא־יֹסִ֖יף לִרְצֹ֣ות עֹֽוד׃** | Shall my lord reject [me] for eternity and never be pleased again? |
| *1 Chr 17:14* | definite | **וְהַֽעֲמַדְתִּ֛יהוּ בְּבֵיתִ֥י וּבְמַלְכוּתִ֖י עַד־הָעֹולָ֑ם** | I will establish it in my house and kingdom unto eternity. |
| *2 Chr 6:2* | plural | **וַֽאֲנִ֛י בָּנִ֥יתִי בֵית־זְבֻ֖ל לָ֑ךְ וּמָכֹ֥ון לְשִׁבְתְּךָ֖ עֹולָמִֽים׃** | I built a lofty house for you, and a place for your dwelling for eternity. |

This case shows that even a term with prototypical adverb behavior can be semantically construed as a noun.[[58]](#footnote-58) Similar to **מָחֳרָת** with **מִן**, **עוֹלָם** appears in nearly all cases with either **לְ** or **עַד** expressing 'until' (157/178); 5 of the 18 cases without 'until' contain nominal modifiers. Given that **עוֹלָם** morphs into 'world' in Medieval Hebrew,[[59]](#footnote-59) this may be a case of semantic shift toward nominalization rather than adverbialization.

Even strongly prototypical nouns can be construed as abstract adverbs. **יוֹם** 'day' occurs with nominal modifiers in 1339/1351 (99%) of sampled cases. Yet the cases below demonstrate extensions of **יוֹם** into an adverb.

Table 17: Select adverbialized samples of *יוֹם*

|  |  |  |  |
| --- | --- | --- | --- |
| reference | modifier | sentence | translation |
| *Isa 43:13* | Ø | **גַּם־מִיֹּום֙ אֲנִ֣י ה֔וּא** | Surely from time past I am he. |
| *Ezek 48:35* | Ø | **וְשֵׁם־הָעִ֥יר מִיֹּ֖ום יְהוָ֥ה׀ שָֽׁמָּה׃** | And the name of city from time past shall be 'YHWH is there'. |

These examples closely resemble in meaning and form the construction **מֵעוֹלָם** 'from eternity'.[[60]](#footnote-60)

These data demonstrate that parts of speech are best modeled as gradient tendencies with prototypicality and variability. Any word associated with a given part of speech association might be construed differently, whether due to hold-over behavior from earlier stages of a word's development, or even due to creative purposes. In English, as in BH and all languages, these kinds of shifts happen all the time. Witness the rise of 'to google', a noun > verb construal. These facts cast suspicion on the classical model's attempt to force single category labels at all costs.

This data also shows that aberrations from the expected part of speech behavior cannot be boiled down to phenomena like ellipsis. For instance, Miller-Naudé and Naudé make the case that **הַטוֹב** 'the good' should be read as elliptical, with **טוֹב** actually modifying a "null noun."[[61]](#footnote-61) The numerous cases reviewed here indicate that such solutions fall short of the scope and scale of variability present in, at minimum, time adverbials.

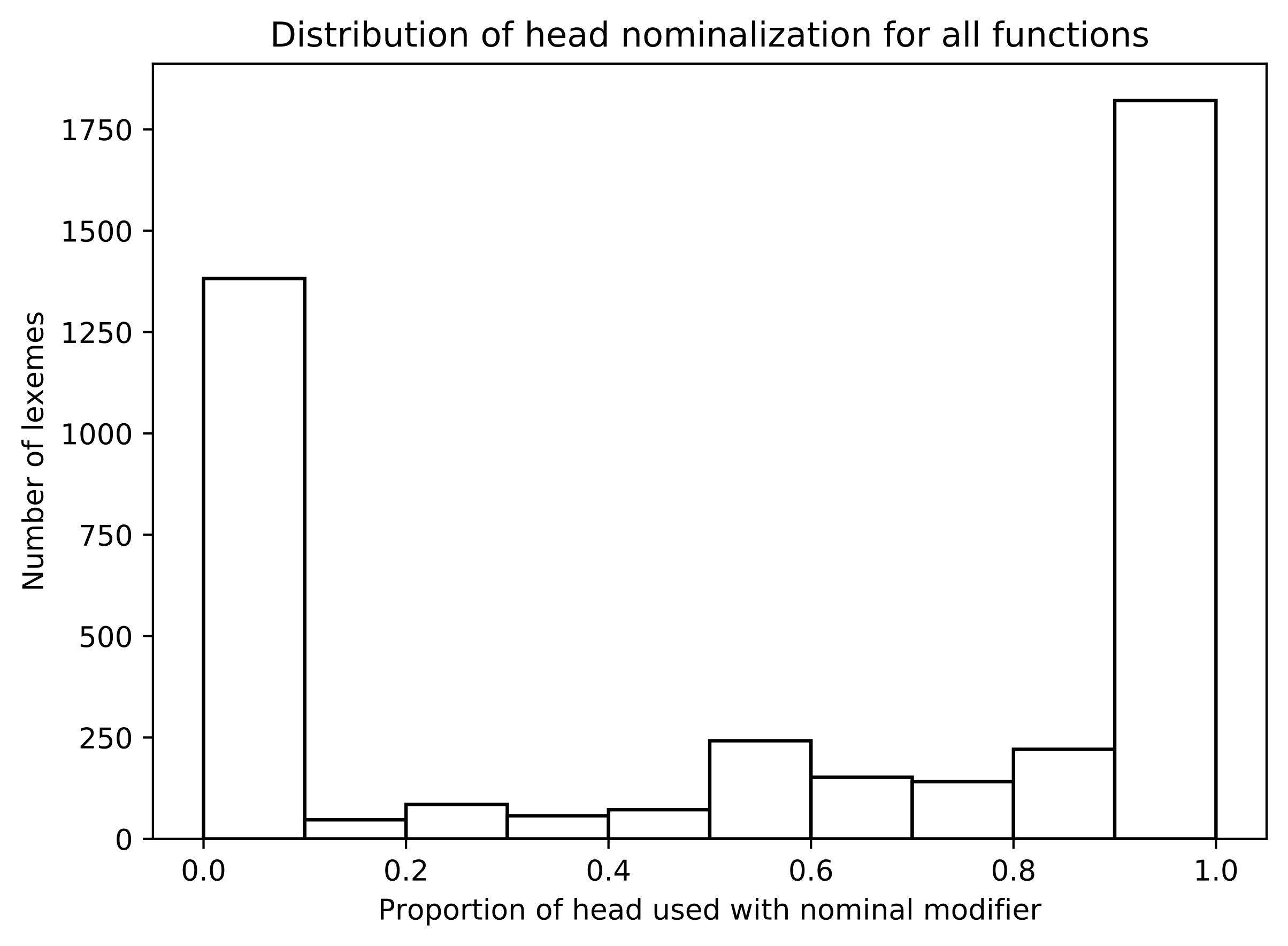
**Head Modifier Collocations**

Up till now, we have examined the way various head modifiers can be used to distinguish adverb and nominal constructions in time adverbials. The next question is how do nominal modifiers uniquely contribute to the semantics of time? Plural morphemes, definite articles, constructs, etc. are sometimes portrayed in grammars as arbitrary, movable components of noun phrases. They are likened, to use an analogy, to a kind of Lego brick. However, quantitative methods reveal subtle differences in the way modifiers are deployed across various phrase functions. Some functions predict more strongly certain modifiers. Modifiers thus resemble magnets more than blocks, with various attractions and repulsions to various roles.

In order to demonstrate these relationships, this section lays out an experiment to test associations between modifiers and functions with a statistical test called ΔP. ΔP has been used in psycholinguistics to formally test associative learning, that is, learning based on a supplied cue and a conditioned response.[[62]](#footnote-62) The more predictive a cue is for a response, the more associated the two categories. ΔP is a unidirectional measure, meaning association is measured with respect to the cue. It is also contingency-based, meaning that the strength of an association is measured over against the strength of other associations and the sample size.[[63]](#footnote-63) ΔP outputs a decimal ratio ranging from -1 to 1.[[64]](#footnote-64) -1 means the response is 100% less likely given the cue (i.e. response is repelled), whereas 1 represents 100% more likely (attracted).[[65]](#footnote-65) For this experiment, each of the 6 phrase functions are modeled as cues with 10 modifiers as responses; the analysis tests how strongly a given function predicts a given modifier.

Before running the test, we first need to isolate a dataset of suitable noun-type samples across 73,120 phrases. This cannot be done manually, and as seen in the previous section head words exhibit a range of behavior. Thus, lexicon parts of speech labels are insufficient. Instead, we will use the statistical behavior of head words to isolate a useful sample. To get an idea of the range of nominal behavior in the dataset, a count is made per every head lexeme of every instance where a nominal modifier occurs.[[66]](#footnote-66) The counts are then normalized to the total frequency of the lexeme in the initial sample. The result is a ratio, which tells what proportion of a given head's total uses appears with a nominal modifier. The range of resulting values is illustrated in the plot below.

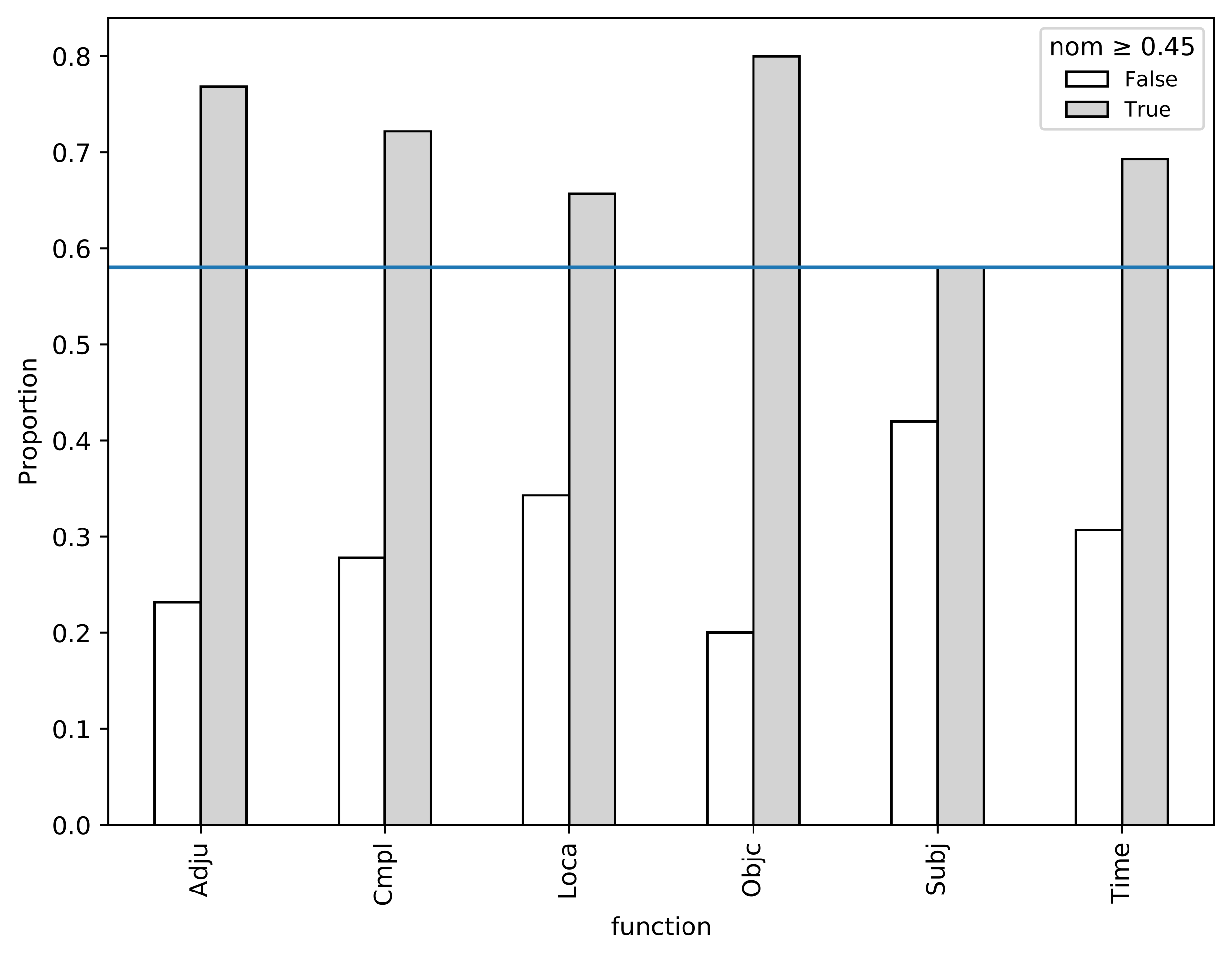
Figure 8: Head lexeme nominal modifier tendencies (N=73,120)



We see, for example, that >1,750 lexemes occur with between 90-100% nominal modification. As seen in the part of speech analysis, a broadly binary division can be seen, with some ambiguity in the center. Earlier, modifier tendencies were measured with sample sizes of at least 5; whereas here there is no minimum. This appears to be why more terms straddle the middle.[[67]](#footnote-67) Given that the majority of nominal uses appear to occur with 50% and greater modification, I selected a cut-off point at 45%, giving a little flexibility. Thus, the sample includes phrases headed by those lexemes with a nominalization ratio of ≥0.45. The resulting dataset contains 50,395 samples (difference of 22,725).

The effects of these selection criteria per function are illustrated in the graph below.

Figure 9: Proportions of nominal phrases by phrase function



Note that the Subj function is most heavily affected. This is due to a high frequency of proper names (e.g. **יְהֹוָה** N=2,169) and pronouns (e.g. **הוּא** N=904). The same is true of Loca (e.g. **שָׁם** N=309; יְרוּשָׁלִַם N=50, etc.).[[68]](#footnote-68) Despite these effects, all functions retain ≥58% of their occurrences, resulting in a large sample of nominal phrases.

For each phrase in the sample, a count of 10 modifier types are made, including all the modifiers from the part of speech experiments as well as the addition of a prepositional modifier. The counts are organized by function. Another table also tallies counts by genre, to control for various genre profiles. The table below contains raw counts for the whole sample.

Table 18: Modifier collocation frequencies for nominal phrases (all genres)[[69]](#footnote-69)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| function | cardinal | definite | demonstrative | genitive | ordinal | plural | preposition | qualitative | suffix | ø |
| *Adju* | 354 | 1175 | 76 | 1530 | 10 | 1372 | 4967 | 291 | 1395 | 652 |
| *Cmpl* | 568 | 3775 | 227 | 3532 | 9 | 3220 | 11877 | 431 | 3684 | 892 |
| *Loca* | 30 | 584 | 32 | 506 | 2 | 244 | 1277 | 69 | 196 | 58 |
| *Objc* | 628 | 1908 | 278 | 2873 | 25 | 4062 | 4249 | 599 | 4307 | 2392 |
| *Subj* | 631 | 3104 | 185 | 3759 | 25 | 5630 | 76 | 756 | 3804 | 1873 |
| *Time* | 411 | 1326 | 491 | 415 | 162 | 562 | 1578 | 209 | 102 | 101 |

The ΔP tests are run on the whole dataset, as well as per genre. The results are illustrated in Figure 10 and Figure 11. To aid interpretation, the resulting data tables are visualized as a heatmap. A heatmap color-codes the various intersecting datapoints with a shade of red, indicating higher attraction, or blue, indicating higher repulsion. The functions along the y-axis are the cues (predictors) whereas the modifiers on the x-axis are the responses.[[70]](#footnote-70) A PCA analysis similar to the one used to analyze Time heads is also shown, which clusters the functions based on their modifier predictions across all genres.

Figure 10: ΔP contingency tests applied to function head modifiers with PCA clustering

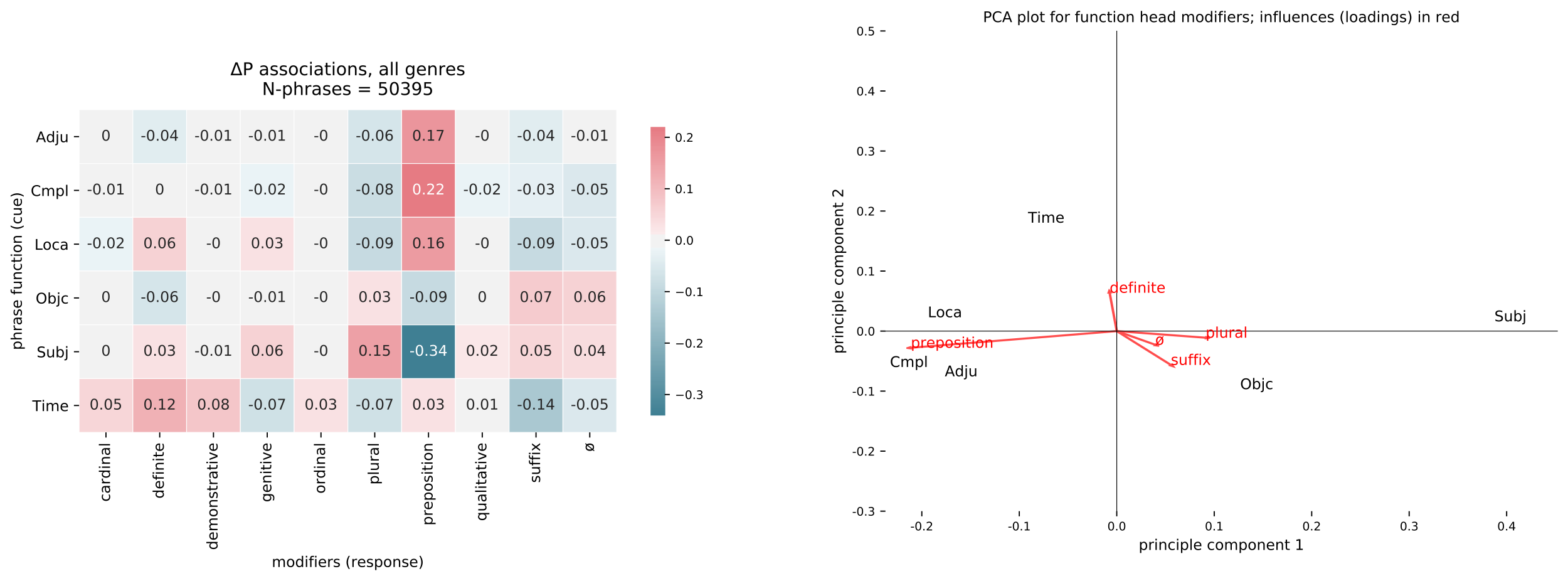
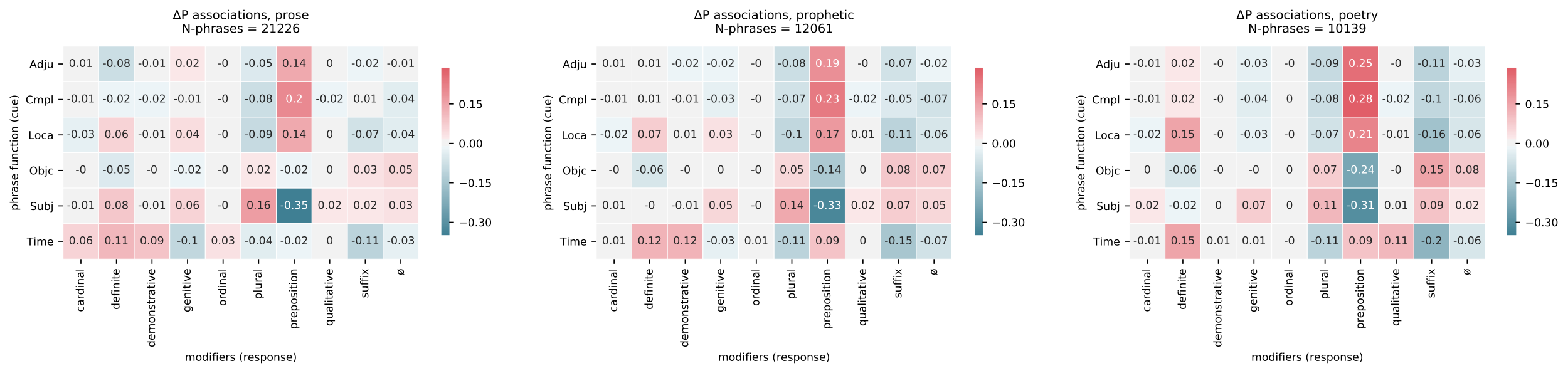


Figure 11: ΔP contingency tests by primary genres

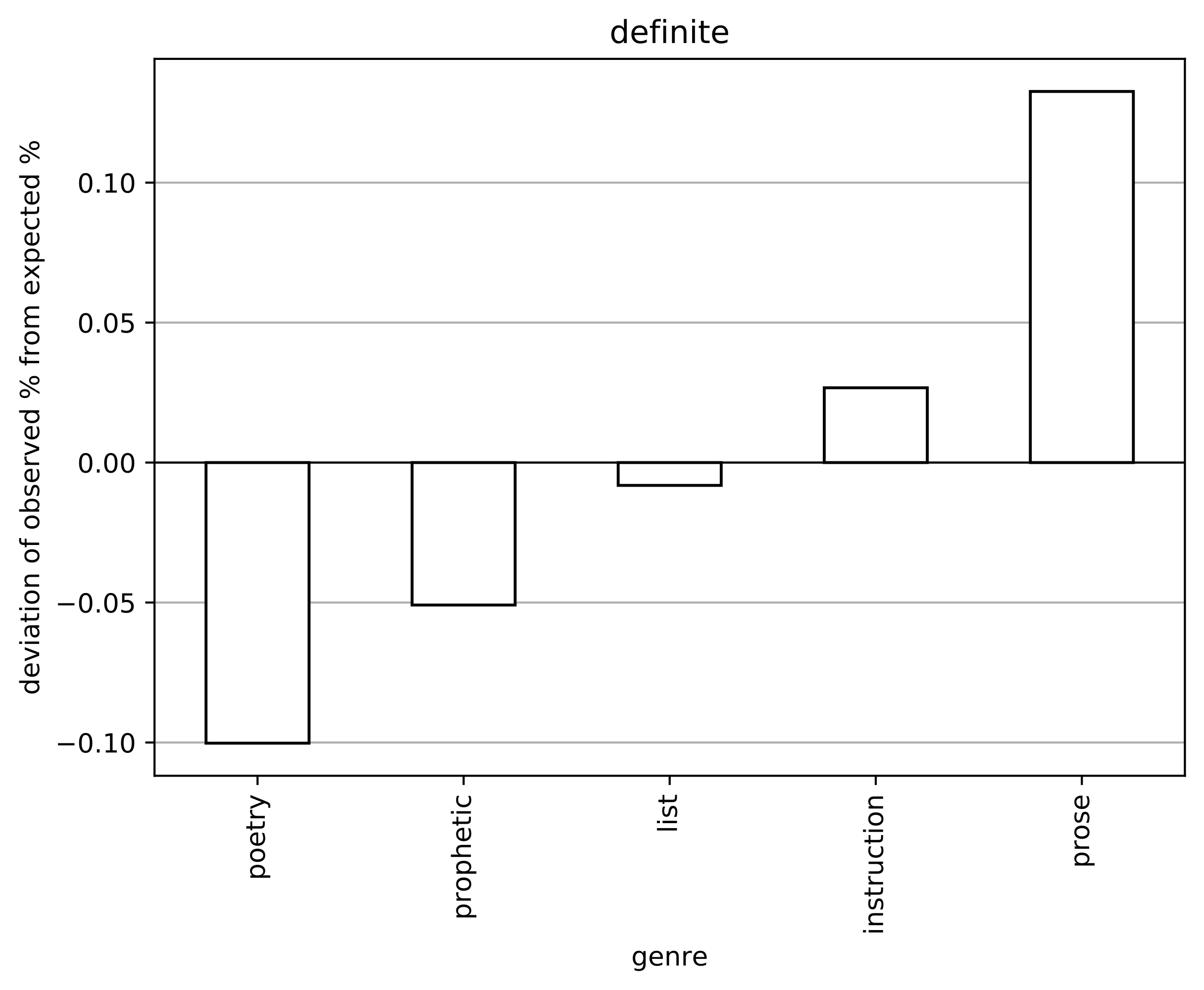


Several noteworthy trends can be seen for Time. For each genre, and in aggregate, Time is most predictive for definiteness (+12% across all genres). This is surprising, since this association is not discussed in the grammars reviewed by this author.[[71]](#footnote-71) The reason for this relationship will be investigated further below. Time also possesses a mixed preference for prepositions; it is weak in prose (-2%) though stronger in prophecy and poetry (+9%). This medial position contributes to its slightly closer placement to Objc in the PCA plot. Time predicts demonstratives (+8%, e.g. **הַיוֹם הַהוּא**), cardinals (+5%, e.g. \*), and ordinals (+3%, e.g. \*). Time has a negative prediction for suffix modification (-14%), plural endings (-7%), and genitive (construct) modifications (-7%). Time does occur with these modifiers, but they are relatively underrepresented; the Subj role, on the other hand, predicts the plural (+15%) as well as the suffix (+5%), a profile it shares with Objc.

What can these tendencies tell about time adverbial semantics? The analysis progresses down two trails. The first follows the intriguing higher preference for definiteness in Time. Rather than a simple coincidence, this tendency appears to reflect a semantic association between nominal time expression and non-deictic anchoring. The second path leads toward the weaker association with prepositions, a behavior predicted by the adverbial accusative lineage present in time indicators. The other attracted modifiers, i.e. cardinals, demonstratives, ordinals, are necessarily handled under these two lines of inquiry.

Time's strong prediction for definiteness across all contexts is a profile it shares with Loca (+6%). Before any functional reasons can be considered for this relationship, it is necessary to first exclude other potential explanations. One alternative explanation is that Time and Loca's stronger representation in prose (see \*ADD PLOT) correlates with a higher representation of definiteness in prose. This possibility is excluded by the fact that definiteness actually increases in poetry (Figure 11), despite a corresponding decrease in Subj and Objc. Furthermore, this tendency actually trends opposite to the overall distribution of definiteness by genre, which is 10% underrepresented in poetry:

Figure 12: Deviation of observed proportions from expected proportions for definiteness by genre[[72]](#footnote-72)



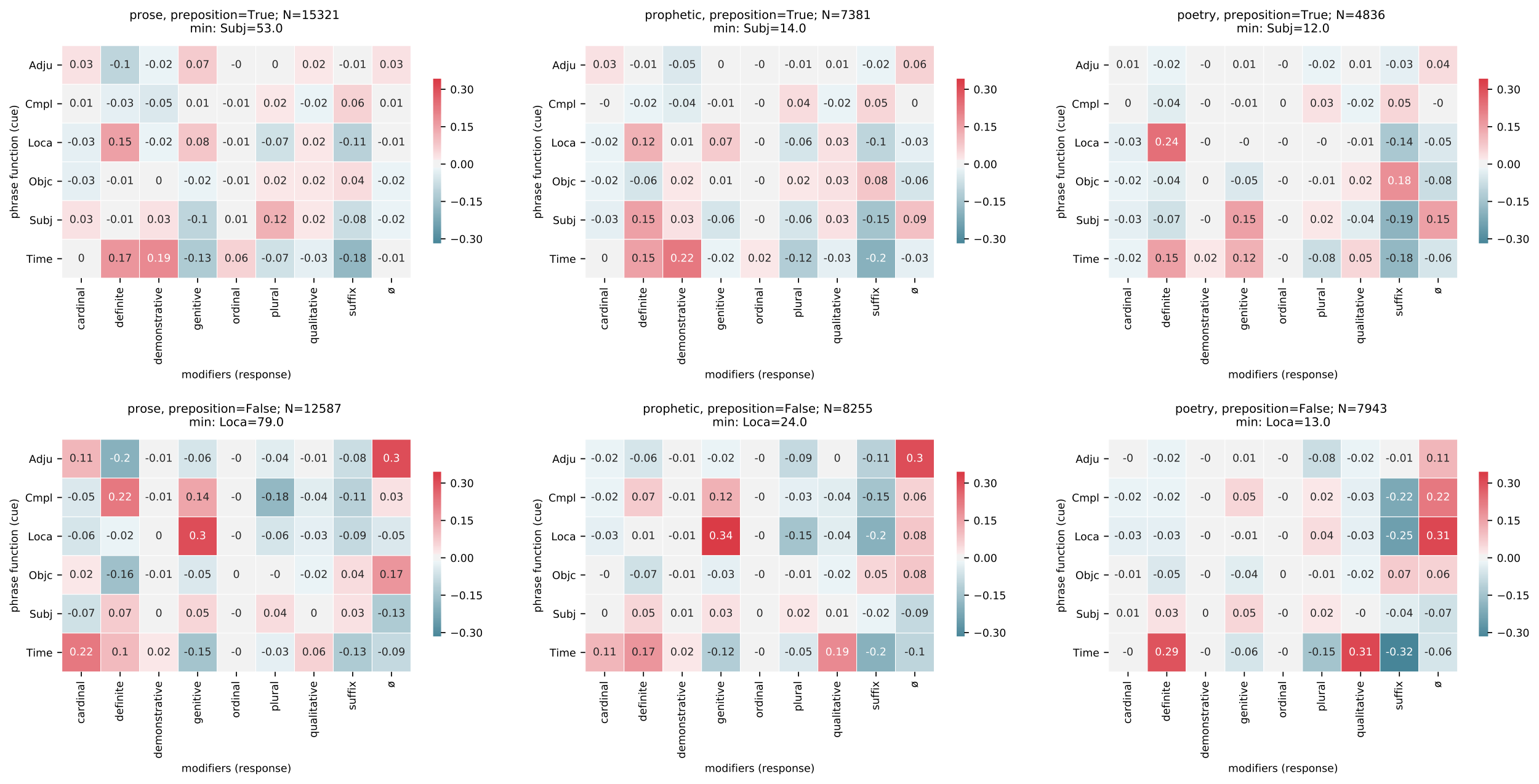
Another possibility is that the higher degree of definiteness can be attributed to Tiberian interference. Gesenius argues that the ubiquitous use of a vocalized article with בְּ + יוֹם was evidence of "textual corruption" by the Masoretes.[[73]](#footnote-73) Barr traces a more detailed argument:

Let it be stated as a useful generalization that a number of fairly common nouns appear in poetry without article, but when combined with the prepositions *b*, *k, l* they are found to be very largely with article. This fact in itself at once makes it very unlikely that determination is the basis for the presence of such articles.[[74]](#footnote-74)

He provides the examples of the words יָם 'sea', שָׁמָיִם 'heavens', and אֶרֶץ 'land' as nouns that are frequently found indefinite in non-prepositional phrases, but definite when combined with a preposition and vocalized article.[[75]](#footnote-75) This discrepancy leads Barr to conclude that these cases may have been "created later by the reading tradition."[[76]](#footnote-76) More recently, Bekins has likewise called into question the originality of such articles.[[77]](#footnote-77)

In order to examine the possibility of Tiberian influence on definiteness, another set of tests are run with and without the preposition across the three main genres (Figure 13). The additional stipulation of "with/without preposition" causes certain functions to have a much lower overall frequency; thus, the lowest observed frequency by function is noted by "min" in the titles.

Figure 13: ΔP tests for modifiers by preposition status and genre



These tests show that Time maintains a predictive status for definiteness even when prepositions are removed (bottom row). In fact, the prediction is generally much higher without the preposition (e.g. +29% for poetry). This likely precludes the possibility of Tiberian influence for Time. However, Loca, which also preferred definiteness, loses its prediction in non-prepositional phrases. Might this be the work of the Masoretes?

There are three reasons to withhold judgment on Loca and definiteness. First, the stipulation of non-prepositional phrases has caused Loca to become significantly underrepresented (as low as N=13 in poetry). This is due to the fact that locative function is intertwined with prepositions in BH.[[78]](#footnote-78) Second, the stipulation has also shifted Loca's prediction from definiteness to genitive in prose and prophecy (+30%, +34%).[[79]](#footnote-79) So it has simply highlighted a different construction, which is irrelevant for the issue of the definite article. Third, there appears to be a convincing functional explanation of why Loca and Time would share this profile, to be explained further.

It is widely recognized that locative adverbials in world languages serve as sources for locating time adverbials.[[80]](#footnote-80) The common definiteness profile of Time and Loca aligns with this hypothesis in a surprising way. Allen and Hill recognized that the definite article (in languages that have them) plays a unique role in location and time phrases.[[81]](#footnote-81) Prototypical adverbs like "tomorrow" or "yesterday" are typically zero marked (Ø) and thus anchored to the speaker's vantage point (i.e. deictic).[[82]](#footnote-82) However, the deployment of a definite article causes the position to be anchored to a different vantage point that is in focus within the discourse.[[83]](#footnote-83) They label these "coding locus" (shared context) and "predicated locus" (external context). They also provide the following illuminative examples:

\*[Two weeks ago Frank...\*Include diagram here]

In the first sentence, "next Monday" is unmarked; thus, "the temporal location of the participants is used as the controlling point in establishing a referent." On the other hand, "the next Monday" portrays the position with respect to Frank's speech instead of the immediate situation shared by the speaker and recipient.[[84]](#footnote-84) Allen and Hill provide a cognitive motivation for this dynamic: zero-marking corresponds with the default here-and-now, whereas definite marking expresses "'there' or 'then'".[[85]](#footnote-85) Consequently, the indication of definiteness sits on a similar axis as the marking of verb tense.[[86]](#footnote-86)

To understand how this dynamic functions in BH Time and Loca, it is helpful to look at the head words that are frequently definite in the sample. The simple construction [preposition + article + head] accounts for 86% of all definite Locas and 25% of all definite Times.[[87]](#footnote-87) The head-word counts for this construction are show below.

Table 19: Head word frequencies in [preposition + article + head] for Time and Loca

|  |  |  |  |
| --- | --- | --- | --- |
| Time head | freq | Loca head | freq |
| **בֹּקֶר** | 100 | **אֶרֶץ** | 64 |
| **עֶרֶב** | 79 | **מִדְבָּר** | 57 |
| **יֹום** | 44 | **שָׂדֶה** | 26 |
| **לַיְלָה** | 23 | **שֶׁמֶשׁ** | 25 |
| **רִאשֹׁון** | 15 | **הַר** | 24 |
| **צָהֳרַיִם** | 12 | **שָׁמַיִם** | 20 |
| **שָׁנָה** | 5 | **דֶּרֶךְ** | 19 |
| **מֹועֵד** | 5 | **עִיר** | 18 |
| **עֵת** | 5 | **בַּיִת** | 18 |
| **חֹדֶשׁ** | 4 | **מִזְבֵּחַ** | 14 |

Both the Time and Loca heads are semantically very general terms. The Time words largely correspond with solar or calendrical cycles; the Loca words correspond with common natural or cultural landmarks. The Time heads align with what Haspelmath calls "canonical times," which are pervasive in world languages:

The major cyclic events of the human natural environment on earth have probably always served as the main means of locating and measuring other situations: in particular, the alternation of light and dark, changes in the shape of the moon, and changes in the path of the sun across the sky (accompanied by marked climatic differences).[[88]](#footnote-88)

The definite article used with these terms correspond with the so-called "global use," which assumes "identifiability" of an item based on general world knowledge.[[89]](#footnote-89) An exploratory manual tagging by this author of all 838 [preposition + article + head] patterns for Time and Loca shows indeed that 85% of Time and 54% of Loca represent a global use.[[90]](#footnote-90) Thus, the definite article preference by Time, and likely for Loca too, is motivated by the need to relate a position in time or space to a known landmark.[[91]](#footnote-91) Whether Masoretic influence factors in to the definiteness of Loca remains an open question, but one that should not be addressed apart from the functional concerns outlined here.[[92]](#footnote-92) Time, at any rate, does not appear to be affected. \*See Waltke §39.3.1h who distinguishes "deictics and independents"

The role of definite modifiers as referential anchors sheds light on the other attracted modifiers, namely, demonstratives, ordinals, and genitives. Namely, non-quantifying modifiers provide landmarks to which a given time word might be anchored.[[93]](#footnote-93) A few representative examples will suffice.

\* put examples of demonstratives, ordinals, and genitives as a table with short descriptions

The last line of inquiry, encompassing the final group of modifiers, is the weak prediction of Time for prepositional modification. The tendency is especially marked given the strong preposition prediction witnessed with Adju, Loca, and Cmpl. A sizeable proportion of nominal Times (XX%) appears without a preposition. The tendency causes Time to be placed closer to Objc in the PCA plot (Figure 10). These are nearly all\* quantified phrases that express a span of time:

\* provide examples here

This matches an old hypothesis in Hebrew studies that durational time adverbials are genealogically related to the proto-Semitic accusative case.[[94]](#footnote-94) Haspelmath independently reaches a similar conclusion in his survey of time adverbials in world languages. He notes that languages with case frequently utilize the accusative with what he calls "atelic extent" time adverbials.[[95]](#footnote-95) These data provide additional support for both Haspelmath's and BH grammarians' hypothesis of an accusative noun source for time adverbials.

**Conclusions**

Using a cognitive-statistical method, this article has examined the collocational tendencies of a large sample of phrases marked for adverbial time function (labeled 'Time') as well as adverbial location, other adverbials, complements, objects, and subjects. The article experimented with a data-driven approach to parts-of-speech for Time heads. A broad division between 'adverb' and 'nominal' classes is confirmed, but with a good degree of variability in between. The data suggests gradient categories better capture the behavior of BH words. Finally, the analysis showed how various nominal modifiers such as the definite article or quantifiers are statistically associated with particular functions. Definiteness was found to be particularly associated with Time due to its use for anchoring time references to well-known entities. The possibility of Tiberian interference in definite prepositional phrases is considered and excluded as a factor for Time phrases. Finally, the analysis finds a lower prediction for prepositions than observed with other adjunctive arguments. This tendency reflects the adverbial accusative lineage present in time adverbials. These modifier data reflect the degree to which constructions in language are not neutral movable pieces, but rather interrelated parts of a complex, interconnected network.

1. See, though, the important work of Van der Merwe (1997). There have also been numerous studies dedicated to individual lexemes or to temporal clauses (כִּי 'for/when', עַתָּה 'now'). There have been two large exegetical studies on time phrases (Brin and DeVries). This study focuses instead on time adverbials as a broader linguistic phenomenon. Christo H.J. Van der Merwe, “Reconsidering Biblical Hebrew Temporal Expressions,” *ZAH* 10, no. 1 (1997): 42–62; James Barr, *Biblical Words for Time* (London: SCM Press, 2005); Alviero Niccacci, “Temporal Clause: Biblical Hebrew,” ed. Geoffrey Khan, *Encyclopedia of Hebrew Language and Linguistics* (Koninklijke Brill NV, 2013), https://doi.org/10.1163/2212-4241\_ehll\_EHLL\_COM\_00000598; Alexey Lyavdansky, “Temporal Deictic Adverbs as Discourse Markers in Hebrew, Aramaic and Akkadian,” *Journal of Language Relationship* 3 (2010): 22–42; Christian Locatell, “Grammatical Polysemy in the Hebrew Bible: A Cognitive Linguistic Approach to כי” (PhD Dissertation, Stellenbosch, South Africa, University of Stellenbosch, 2017); 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); Simon J. DeVries, *Yesterday, Today and Tomorrow: Time and History in the Old Testament* (USA: Eerdmans, 1975). [↑](#footnote-ref-1)
2. Van der Merwe asked similar questions in 1997: "Is there any difference in the syntax of non-temporal adjuncts and temporal adjuncts? Which BH constructions can function as temporal adjuncts? Apart from the above-mentioned semantic classes of temporal position, duration and frequency, are there other semantic classes or subclasses to be identified among temporal adjuncts?" Van der Merwe, “Reconsidering Temporal Expressions,” 49. [↑](#footnote-ref-2)
3. 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; A. Dean Forbes, “Distributionally-Inferred Word and Form Classes in the Hebrew Lexicon: Known by the Company They Keep,” in *Foundations for Syriac Lexicography II*, ed. Peter J. Williams (Piscataway, NJ: Gorgias Press, 2009), 1–34. [↑](#footnote-ref-3)
4. Barbara C. Scholz, Francis Jeffry Pelletier, and Geoffrey K. Pullum, “Philosophy of Linguistics,” in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, 2011, https://plato.stanford.edu/archives/win2016/entries/linguistics/; Dirk Geeraerts, *Theories of Lexical Semantics* (Oxford ; New York: Oxford University Press, 2010). [↑](#footnote-ref-4)
5. Evelina Fedorenko et al., “Lack of Selectivity for Syntax Relative to Word Meanings throughout the Language Network,” *Cognition* 203 (October 2020): 104348, https://doi.org/10.1016/j.cognition.2020.104348; Jacob Devlin et al., “BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding,” *ArXiv:1810.04805 [Cs]*, May 24, 2019, http://arxiv.org/abs/1810.04805; Florent Perek and Adele E. Goldberg, “Linguistic Generalization on the Basis of Function and Constraints on the Basis of Statistical Preemption,” *Cognition* 168 (November 2017): 276–93, https://doi.org/10.1016/j.cognition.2017.06.019. [↑](#footnote-ref-5)
6. Martin Haspelmath, “Pre-Established Categories Don’t Exist: Consequences for Language Description and Typology,” *Linguistic Typology* 11 (2007): 119–32. [↑](#footnote-ref-6)
7. 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), 380–81. [↑](#footnote-ref-7)
8. \*get online ref Dirk Geeraerts and Hubert Cuyckens, “Introducing Cognitive Linguistics,” in *The Oxford Handbook of Cognitive Linguistics*, ed. Dirk Geeraerts and Hubert Cuyckens (Oxford: Oxford University Press, 2012), §2. [↑](#footnote-ref-8)
9. I agree with Ellis et al.: "We take the Saussurian (1916) view that the units of language are constructions—form-meaning mappings, conventionalized in the speech community, and entrenched as language knowledge in the learner’s mind." 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. For experiential categories see George Lakoff, *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind* (Chicago: University of Chicago Press, 1987). [↑](#footnote-ref-9)
10. Lakoff, *Women, Fire, and Dangerous Things*, 41. [↑](#footnote-ref-10)
11. Steven T. Piantadosi, “Zipf’s Word Frequency Law in Natural Language: A Critical Review and Future Directions,” *Psychonomic Bulletin & Review* 21, no. 5 (October 2014): 1112–30, https://doi.org/10.3758/s13423-014-0585-6; Ellis, O’Donnell, and Römer, “Usage-Based Language,” 31–33. [↑](#footnote-ref-11)
12. On similarity-based categorization see Vladimir M. Sloutsky et al., “Conceptual Influences on Induction: A Case for a Late Onset,” *Cognitive Psychology* 82 (November 2015): 1–31, https://doi.org/10.1016/j.cogpsych.2015.08.005. [↑](#footnote-ref-12)
13. 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), 355–80. [↑](#footnote-ref-13)
14. Indeed, as noted above, these mental associations are learned over relationships observed in the world. Language thus represents a re-externalization of semantic input. [↑](#footnote-ref-14)
15. Famously summarized by Firth as "knowing a word by the company it keeps." J.R. Firth, “A Synopsis of Linguistic Theory, 1930–1995,” in *Studies in Linguistic Analysis* (Oxford: Basil Blackwell, 1962), 1–32; R. Xiao, “Collocation,” in *The Cambridge Handbook of English Corpus Linguistics*, ed. D. Biber and R. Reppen (Cambridge: Cambridge University, 2015), 106–24. [↑](#footnote-ref-15)
16. Based on an ongoing analysis by this author: https://github.com/codykingham/noun\_semantics/blob/master/analysis.ipynb [↑](#footnote-ref-16)
17. Anatol Stefanowitsch and Stefan Th. Gries, “Collostructions: Investigating the Interaction of Words and Constructions,” *International Journal of Corpus Linguistics* 8, no. 2 (2003): 209–43. [↑](#footnote-ref-17)
18. See, for instance, Devlin et al., “BERT.” [↑](#footnote-ref-18)
19. See Carson T. Schütze, *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*, Classics in Linguistics 2 (Berlin: Language Science Press, 2016). [↑](#footnote-ref-19)
20. See Dirk Geeraerts, “The Doctor and the Semantician,” in *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*, ed. Dirk Geeraerts and John R. Taylor, Cognitive Linguistics Research 46 (Berlin: De Gruyter Mouton, 2010), 63–78. [↑](#footnote-ref-20)
21. A. Dean Forbes, “Squishes, Clines, and Fuzzy Signs: Mixed and Gradient Categories in the Biblical Hebrew Lexicon,” in *Foundations for Syriac Lexicography: Colloquia of the International Syriac Language Project*, ed. Beryl Turner et al., Perspectives on Syriac Linguistics 1 (Piscataway, NJ: Gorgias Press, 2005), 105–39; Forbes, “Distributionally-Inferred.” [↑](#footnote-ref-21)
22. This is the process known as 'operationalization' and is the "secret" to the success of the natural sciences. Stefanowitsch, “Empirical Cognitive Semantics,” 358–61. [↑](#footnote-ref-22)
23. Dirk Roorda et al., “Biblia Hebraica Stuttgartensia (Amstelodamensis) Documentation,” Documentation, ETCBC Github, April 10, 2019, https://etcbc.github.io/bhsa/. [↑](#footnote-ref-23)
24. 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). [↑](#footnote-ref-24)
25. Reinoud Oosting, “Computer-Assisted Analysis of Old Testament Texts: The Contribution of the WIVU to Old Testament Scholarship,” in *The Present State of Old Testament Studies in the Low Countries: A Collection of Old Testament Studies Published on the Occasion of the Seventy-Fifth Anniversary of the Oudtestamentisch Werkgezelschap*, ed. Klaas Spronk, vol. 69, Oudtestamentische Studiën (Leiden: Brill, 2016), 192–209. [↑](#footnote-ref-25)
26. This is done using a custom Python parser which works within the existing phrase segmentation of the BHSA. In some cases, the phrase segmentations are modified where they are deficient. The parser can be viewed in the project repository at https://github.com/CambridgeSemiticsLab/BH\_time\_collocations/blob/master/data/cxs/phrase\_grammar.py [↑](#footnote-ref-26)
27. Anatol Stefanowitsch, *Corpus Linguistics: A Guide to the Methodology*, Textbooks in Language Sciences 7 (Berlin: Language Science Press, 2020), 111–16. [↑](#footnote-ref-27)
28. Stefanowitsch, 113. [↑](#footnote-ref-28)
29. \*archived link [↑](#footnote-ref-29)
30. I used Croft's notion of semantic headship as a general guide. Objects of prepositions are the semantic heads of prepositional phrases. Quantifiers including כֹל 'all' are excluded as heads if they modify a word. William Croft, *Radical Construction Grammar: Syntactic Theory in Typological Perspective* (Oxford ; New York: Oxford University Press, 2001), 257–59. [↑](#footnote-ref-30)
31. A measure called deviation of proportions (Gries 2008) was used to compare the sampled distribution by book against the whole dataset. The most affected books are Deut (3% total deviation), Ps (3%), 2 Chr (2%), Isa (2%), and Job (2%). Stefan Th. Gries, “Dispersions and Adjusted Frequencies in Corpora,” *International Journal of Corpus Linguistics* 13, no. 4 (2008): 415–19. [↑](#footnote-ref-31)
32. Throughout the study I will refer to ratios out of 1, which can be converted to percentages by multiplying them by 100. This is standard practice in statistical reporting. \* [↑](#footnote-ref-32)
33. This function is more heavily affected by the selection criteria than others. This is due to the exclusion of prepositional phrases that contain only a preposition + suffix. Technically the suffix serves as the head of these phrases, but for now I must exclude these cases since suffixes are not modeled as words in the database. [↑](#footnote-ref-33)
34. Stefanowitsch, “Empirical Cognitive Semantics,” 358–61. [↑](#footnote-ref-34)
35. E.g. 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. [↑](#footnote-ref-35)
36. E.g. William Croft, “A Conceptual Framework for Grammatical Categories (or: A Taxonomy of Propositional Acts),” *Journal of Semantics*, no. 7 (1990): 245–79. [↑](#footnote-ref-36)
37. Paolo Ramat, “Adverbial Grammaticalization,” in *The Oxford Handbook of Grammaticalization*, ed. Bernd Heine and Heiko Narrog (Oxford University Press, 2011), https://doi.org/10.1093/oxfordhb/9780199586783.013.0040; Pierluigi Cuzzolin, Putzu Ignazio, and Ramat Paolo, “The Indo-European Adverb in Diachronic and Typological Perspective,” *Indogermanische Forschungen* 111 (2006): 1–38. [↑](#footnote-ref-37)
38. A handful of terms possess adverb suffixes such as ָם as found on **יוֹמָם** 'by day'. Van der Merwe, Naudé, and Kroeze, *BHRG*, 380. [↑](#footnote-ref-38)
39. Dorit Ravid and Yitzhak Shlesinger, “Modern Hebrew Adverbials: Between Syntactic Class and Lexical Category,” in *Between Grammar and Lexicon*, ed. Ellen Contini-Morava and Y. Tobin, Amsterdam Studies in the Theory and History of Linguistic Science, v. 183 (International Cognitive Linguistics Conference, Amsterdam: John Benjamins, 2000), 339–44. [↑](#footnote-ref-39)
40. This includes articles only reflected in the vocalized text. [↑](#footnote-ref-40)
41. In this case the term of interest is the *nomen rectum*, or trailing term in a construct chain, rather than the word morphologically marked as construct. This is the reason the term 'genitive' is used. The BHSA dataset includes constructs without explicit marking. [↑](#footnote-ref-41)
42. The attributive adjunct construction involving the definite article, i.e. הַ + word + הַ + adjunct, is also tagged separately for additional analysis. But since that construction itself involves definite modification, it is kept out of the initial dataset. The simple adjunct construction of word + adjunct is not yet tagged due to existing limitations in the dataset. In the BHSA, these relations are tagged as so-called 'subphrase relations'; but subphrase relations suffer from a number of problems and inadequacies making them unreliable. In the case of time adverbials, it seems (based on previous analyses) that adjectival elements play a lesser role in the semantics of time. Nevertheless, in the future it would be better to include these modifiers in the dataset. [↑](#footnote-ref-42)
43. An alternative normalization would be a measure of statistical significance as used elsewhere in this study. But in this case, a ratio makes more sense because we want to compare word tendencies to one another rather than find particular distinctives of each individual word. [↑](#footnote-ref-43)
44. Principle component analysis is a dimensionality reduction method which seeks to maximize the variation between the variable vectors (i.e. lists of observed ratios), and thus separate samples along an axis of greatest variation (the first principle component) and second greatest variation (the second). The first principle component can be used as an x-axis value and the second as the y-axis value, yielding a graph of samples which are placed nearer or farther to one another based on their features. Vladimir Rokhlin, Arthur Szlam, and Mark Tygert, “A Randomized Algorithm for Principal Component Analysis,” *SIAM Journal on Matrix Analysis and Applications* 31, no. 3 (January 2010): 1100–1124, https://doi.org/10.1137/080736417; F. Pedregosa et al., “Scikit-Learn: Machine Learning in Python,” *Journal of Machine Learning Research* 12 (2011): 2825–2830. [↑](#footnote-ref-44)
45. The lexicon parts of speech values come from the BHSA's 'sp' feature, which is derived from KBL2. The abbreviated values are substantive, interrogative, adjective, and adverb, respectively. Ludwig Köhler and Walter Baumgartner, *Supplementum Ad Lexicon in Veteris Testamenti Libros* (Leiden: Brill, 1958). [↑](#footnote-ref-45)
46. These are the loading scores, which are the following the discussion of ttnphns (https://stats.stackexchange.com/users/3277/ttnphns), “Loadings vs Eigenvectors in PCA: When to Use One or Another?,” Stack Exchange, March 29, 2015, https://stats.stackexchange.com/q/143949. [↑](#footnote-ref-46)
47. The loading score includes direction from the eigen vector and magnitude from the eigen value. The Ø modifier has a PC1 (x-axis) of -4.8 and PC2 (y-axis) of 3.9. The line is drawn from origin to this point. [↑](#footnote-ref-47)
48. E.g. Gen 7:11; Deut 3:14; Josh 20:6; 1 Sam 7:6; 2 Sam 3:38; 2 Kgs 20:1; Isa 28:5; Ezek 39:11; Zech 13:4; 2 Chr 35:16. [↑](#footnote-ref-48)
49. E.g. Isa 47:12, 15; Judg 16:30, 2 Sam 18:18. [↑](#footnote-ref-49)
50. In Modern Hebrew the distinction between adjectives and adverbs is blurry. A similar dynamic appears in BH. Ravid and Shlesinger, “Modern Hebrew Adverbials,” 346. [↑](#footnote-ref-50)
51. Compared to words with physical manifestations like יוֹם 'day' or עֶרֶב 'evening'. [↑](#footnote-ref-51)
52. Jer 20:3, 2 Kgs 8:15, 1 Sam 31:8, 1 Sam 18:10, 1 Sam 11:11, Judg 21:4, Judg 9:42, 1 Chr 10:8, Num 17:23, Exod 32:30, Exod 18:13. [↑](#footnote-ref-52)
53. See the subsequent section regarding the ongoing discussion about the status of vocalized articles in the Tiberian tradition. [↑](#footnote-ref-53)
54. Thanks to Chip Hardy for pointing out these possibilities in relation to another similar case, לְעוֹלָם. [↑](#footnote-ref-54)
55. N.B. that this table contains a higher marginal total than the previous since modifiers can co-occur. [↑](#footnote-ref-55)
56. The closest would be בֹּקֶר with 18/119 (15%) null modified uses, a substantial difference from the 35% of לַיְלָה. יוֹם has 0.08% Ø; עֶרֶב has 3.5% Ø. [↑](#footnote-ref-56)
57. Thanks to Chip Hardy for pointing this out. \*literature backing this up [↑](#footnote-ref-57)
58. Another case is נֵצַח 'everlasting' with 96% null modification (27/28) and 1 nominal construal in Isa 34:10: **לְנֵ֣צַח נְצָחִ֔ים אֵ֥ין עֹבֵ֖ר בָּֽהּ** [↑](#footnote-ref-58)
59. Angel Sáenz-Badillos, *A History of the Hebrew Language*, Reprinted (Cambridge: Cambridge University Press, 1997), 200. [↑](#footnote-ref-59)
60. Josh 24:2; Isa 42:14; Isa 57:11; Isa 63:19; Isa 64:3; Jer 2:20; Prov 8:23. [↑](#footnote-ref-60)
61. Miller-Naudé and Naudé, “A Re-Examination,” 288–303. [↑](#footnote-ref-61)
62. Nick C. Ellis, “Language Acquisition as Rational Contingency Learning,” *Applied Linguistics* 27, no. 1 (March 1, 2006): 10–12, https://doi.org/10.1093/applin/ami038. [↑](#footnote-ref-62)
63. Explain parts of contingency table, a, b, c, d, etc.\* [↑](#footnote-ref-63)
64. Ellis, “Language Acquisition,” 11. [↑](#footnote-ref-64)
65. That is, the probability of a response given the cue minus the probability of the response without the cue. Ellis, 11. [↑](#footnote-ref-65)
66. A True/False value is stored on every instance in the sample for whether a word contains a nominal modifier; the value is 'True' if there is any of the following: plural ending, \* [↑](#footnote-ref-66)
67. A lexeme that occurs twice, for instance, might have 1 nominal modifier and 1 Ø, yielding a ratio of 0.5 (50%). [↑](#footnote-ref-67)
68. This is also a vindication of using a statistical method rather than a lexicon label since certain conceptual 'nouns' (e.g. proper names) may not have proto-typical noun behavior. [↑](#footnote-ref-68)
69. Note that the marginal frequency for the whole table is greater than the sample size since modifiers may co-occur in the same phrase. [↑](#footnote-ref-69)
70. Typically the x-axis would contain the independent variables. But in this case I found the values to be more readable from y to x-axis, indicating "y cues x". [↑](#footnote-ref-70)
71. \*Give literature; Gesenius mentions the frequent use of הַ together with יוֹם as well as its ubiquitous use with בּ and יוֹם §126w; the demonstrative use of the article with time adverbials is mentioned frequently\* (Gesenius §126b) [↑](#footnote-ref-71)
72. For an explanation of this calculation, see the first section on distribution. [↑](#footnote-ref-72)
73. Gesenius, Kautsch, and Cowley, *GKC*, §126w. [↑](#footnote-ref-73)
74. James Barr, “‘Determination’ and the Definite Article in Biblical Hebrew,” *Journal of Semitic Studies* 34, no. 2 (1989): 327. [↑](#footnote-ref-74)
75. Barr, 326. [↑](#footnote-ref-75)
76. Barr, 327. [↑](#footnote-ref-76)
77. \*Bekins 2016 [↑](#footnote-ref-77)
78. Humphrey Hardy and Samuel Boyd, “Hebrew Adverbialization, Aramaic Language Contact, and Mpny ʾšr in Exodus 19:18,” in *Semitic Languages in Contact*, ed. Aaron Butts (Brill, 2015), https://doi.org/10.1163/9789004300156. [↑](#footnote-ref-78)
79. The Loca sample size N=13 in poetry means there is not enough data to make a definitive assessment for that genre. [↑](#footnote-ref-79)
80. 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-80)
81. Robert L. Allen and Clifford A. Hill, “Contrast between ø and *the* in Spatial and Temporal Predication,” *Lingua* 48, no. 2–3 (June 1979): 123–46, https://doi.org/10.1016/0024-3841(79)90002-0. [↑](#footnote-ref-81)
82. Haspelmath, *From Space to Time*, 43. [↑](#footnote-ref-82)
83. Allen and Hill, “Contrast,” 126–27. [↑](#footnote-ref-83)
84. Allen and Hill, 135. [↑](#footnote-ref-84)
85. Allen and Hill, 140–41. [↑](#footnote-ref-85)
86. Allen and Hill, 138. See also Croft's remarks: "...there does seem to be a semantic parallel between the category of definiteness in nouns...and modal or modality related categories in verbs and complex sentence constructions such as world-creating predicates and conditional constructions." Croft, “A Conceptual Framework,” 265. [↑](#footnote-ref-86)
87. Though only 25% with Time, it is the second-most frequent pattern and it also appears as a sub-construction within the most common pattern, definite attributive apposition, e.g. בַיוֹם הַהוּא. [↑](#footnote-ref-87)
88. Haspelmath, *From Space to Time*, 25. [↑](#footnote-ref-88)
89. Peter Bekins, “Non-Prototypical Uses of the Definite Article in Biblical Hebrew,” *Journal of Semitic Studies* 58, no. 2 (October 1, 2013): 227, https://doi.org/10.1093/jss/fgt001. [↑](#footnote-ref-89)
90. While this analysis extends beyond the bounds of this article, the results of the preliminary experiment and the tagging itself can be viewed at \*link. [↑](#footnote-ref-90)
91. For "landmarks" see 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). [↑](#footnote-ref-91)
92. The absence of definite marking does not automatically imply a word is not actually definite. Proper names or well-known entities might be used without definite articles but remain semantically definite by virtue of their recognizability. \*ref. Thus, the issue for Loca remains complicated, related to the way dialectal differences may deploy definite articles. \*Conversation with Ben Kantor. [↑](#footnote-ref-92)
93. 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), 81; Fillmore, “Mini-Grammars of Some Time-When Expressions in English,” 38. [↑](#footnote-ref-93)
94. Hardy, grammars, Matthew Anstey, etc.\* [↑](#footnote-ref-94)
95. English is actually an exception in this case, preferring to use "for" to indicate such durational spans. \*Haspelmath [↑](#footnote-ref-95)