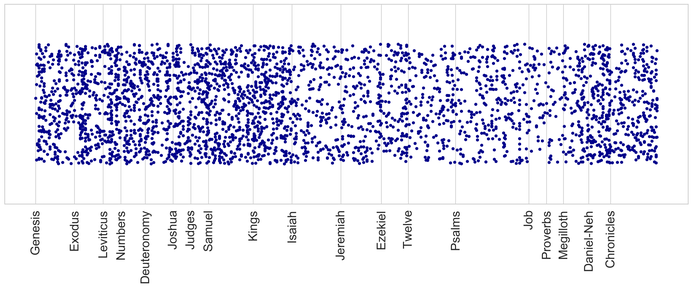
Time Adverbial Collocations in Biblical Hebrew

Cody Kingham, PhD candidate • Geoffrey Khan, supervisor
cak47@cam.ac.uk • github.com/CambridgeSemiticsLab/BH_time_collocations

Introduction

Most research on temporality in Biblical Hebrew focuses on the semantics and pragmatics of the verbal system. Yet time adverbials also contribute key information to temporality and are comparatively understudied. They can anchor verbal event time, frame discourse, or direct pragmatic focus. A time adverbial is defined as a clause constituent which can consist of a word or phrase that modifies the event time of a clause. There is thus great diversity in the forms used to express time adverbial function. This project combines insights from construction grammar (Goldberg 1995) with methods from computational corpus linguistics in order to map the form and function of the main time adverbials in Biblical Hebrew. A usage-based framework allows phrase functions to be analyzed with statistical and empirical methods.

Time Adverbial Distribution and Identity



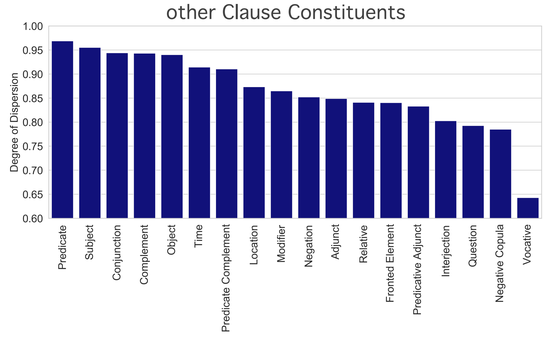
This chart illustrates the distribution of 3,881 time adverbials as selected in the ETCBC Hebrew syntax database (BHSA). The distribution shows higher concentrations in narrative books. The x-axis consists of chapter-by-chapter slices upon which the adverbials are plotted; the y-axis is generated randomly using a "jitter" effect. These adverbials provide a starting point for evaluating the typical forms and functions.

Top 50 Time Adverbial Surface Forms, (dot-separated words, bigger is more frequent)

To get an idea of the main forms, adverbial phrases marked with a function of "TIME" in the BHSA are tokenized by stripping accents and joining words with dots. These are then counted. The most frequent forms illustrate prototypical constructions.

Degree of dispersion calculates how well-distributed an item is across a corpus's parts (Gries 2008). These measures, taken across chapters, reveal Time to be more regular than other adjuncts (Location, Adjunct).

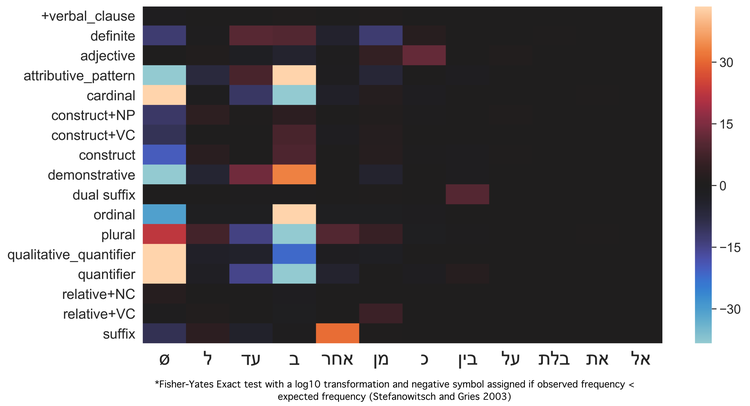
Uniformity of Distribution Compared to other Clause Constituents



Collocations within Time Adverbials

What role do the various components within a time adverbial play for its function? Collocation analysis reveal statistical associations between adverbial constituents, which gives a clue to their semantic functions. An algorithm searches inside and around each time adverbial and counts constituents such as "pl" (plural ending). Raw counts are converted to association scores with a Fisher-Yates Exact test (Stefanowitsch and Gries 2003). A score >1.3 is statistically attracted while < -1.3 is statistically repelled. Below, for each adverbial a count is made of its preposition and its constituent elements. The statistical data is represented as a heat-map.

Preposition and Adverbial Constituent Collocations (red is attraction, blue is repulsion*)



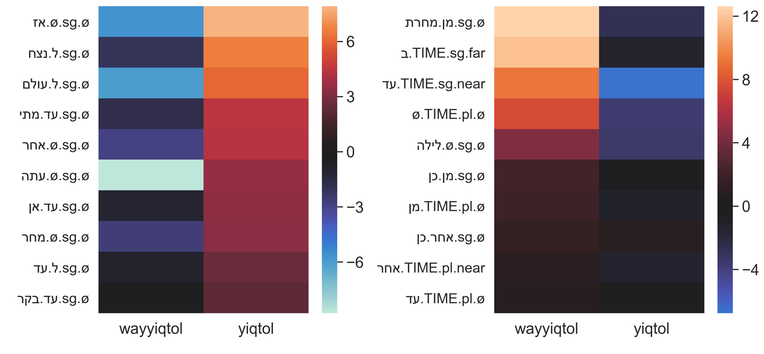
*Fisher-Yates Exact test with a log10 transformation and negative symbol assigned if observed frequency < expected frequency (Stefanowitsch and Gries 2003)

The scores reveal the opposition between preposition-less (ø) time adverbials and those marked with ב. Everywhere ø is red, ב is blue. The ø's preference for quantifiers reveals its function as the durative marker. Note also the subtle opposition between ל and עד. The ל adverbials prefer plurals, while עד prefers demonstratives and definites.

Collocations with Verbs

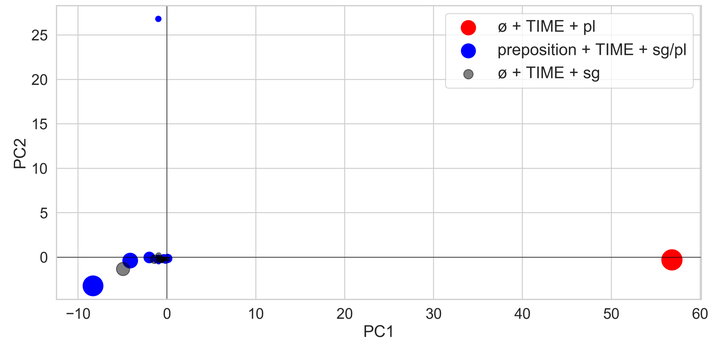
One of the main goals of classifying time adverbials is to examine their interaction with other clause constituents. The interaction of adverbials and verbs may contribute to the ongoing discussion on the verbal system. The tables below show statistical associations (Fisher-Yates Exact tests) between the yiqtol and wayyiqtol tenses with classified time adverbials. "TIME" is substituted in cases where the time word varies often in the given construction. "near" or "far" is substituted for a respective demonstrative.

Yiqtol and Wayyiqtol Time Adverbial Preferences



The yiqtol shows a marked preference for time adverbials with heads that are prototypical "adverbs" while the wayyiqtol prefers noun-based times. In general, the yiqtol prefers unbound durations (e.g. לעולם) while wayyiqtol prefers punctual times (e.g. near demonstratives, ב+TIME).

Opposition between Durative and Punctual Adverbials Based on Verb Lexeme Preference (PCA)



Verbal tense is not the only factor at play when it comes to adverbial / verb collocations. The verbal lexeme also contributes aspectual information ("Aktionsart") that attracts/repels corresponding time adverbials. The chart above shows a PCA clustering method that compares the adverbial + verb lexeme collocations amongst three groups of time adverbials. The PCA algorithm separates unmarked plurals, which is the durative construction, from the rest. This shows an opposition between durative and punctual time adverbials. The separation is based on verb lexemes such as מלך "to reign," which frequently collocate with durative adverbials, for example שתי שנים "for two years" (1 Sam 13:1).