

Agency and lexical decomposition of Biblical Hebrew verbs

Christian Canu Højgaard

Fjellhaug International University College Copenhagen; & Vrije Universiteit, Amsterdam
cch@dbi.edu

Abstract: This paper readdresses Dowty’s classical approach to lexical decomposition of verbs. Dowty’s approach, which is incorporated into Role and Reference Grammar, presupposes a certain intuition for the language in question; an intuition that cannot be taken for granted for Biblical Hebrew. Accordingly, quantitative methods, including “covarying collexeme analysis” and Principal Component Analysis, and new qualitative criteria are explored in this paper in order to propose the internal temporal aspect (also known as *Aktionsart*) of Biblical Hebrew verbs. In particular, the oppositions between stative and active verbs and between causative and non-causative verbs are scrutinized.

1 Introduction

One of the classical components of lexical decomposition of verbs is Dowty’s (1979) famous interrogative questions. The interrogative tests have been incorporated into the RRG theory as binary categories. In canonical RRG, then, there are seven test-questions to determine the *Aktionsart* of any given verb, and the framework has proven useful for many languages (Van Valin 2005, 39).

When working with ancient corpora, things become more complicated. One observation that quickly emerges from applying Dowty’s test-questions is that the answer to those questions intrinsically relies on a certain intuition for the language. In short, the basic assumption of the test-questions is the researcher’s – or his respondents’ – ability to sort out meaningful senses from awkward ones. The obvious unacceptable sentence in (1a) is not deemed unacceptable because of the syntactical construction but because of our making sense of the verb “shatter” which does not accept the adverbial “vigorously” in contrast to “shake” (b) (Van Valin 2005, 36):

- (1) a. *The window shattered vigorously.
b. The house is shaking vigorously.

We cannot assume such an intimate knowledge of ancient languages for which there are no living language users to consult. Although attempts have already been made to classify Hebrew verbs within an RRG framework (Winther-Nielsen 2016), we may need to address the whole question of inherent verbal aspect from a *quantitative* point of view, rather than a *qualitative*.

Accordingly, the aim of this paper is to explore quantitative, statistical methods for disambiguating stative and active verbs. At the end of this paper, I will extend the analysis of Hebrew verbs to include another fundamental aspect, that is, causation.

2 A quantitative approach to classifying stative/active verbs

The principal parameter for distinguishing stative from active verbs is the ability of the verb to combine with dynamic adverbs, e.g. “vigorously”, “gently”, and “powerfully”. However, unlike the traditional qualitative tests, I will apply a statistical test to explore how frequent Hebrew verbs combine with certain dynamic adverbials.

The analysis is conducted with a method called “covarying collexeme analysis” developed by Stefan Gries and Anatol Stefanowitsch (Gries and Stefanowitsch 2004; Stefanowitsch and Gries 2005) within the framework of Construction Grammar. The method is developed to investigate the interrelation between two similar slots of the same construction. In the present case, the two respective slots are the predicate and the complement. For the sake of time, I will not go into a long explanation of the method. It suffices to say that the statistical method is applied in order to compare how often a certain combination of a verb and an adverbial occurs in the same construction, in contrast to how frequent the verb occurs with other adverbials and how often the adverbial occurs with other verbs.

2.1 Corpus and method

The research is carried out on a linguistically annotated corpus of the Hebrew Bible, called the ETCBC database. The database is richly annotated, including important annotations for our purposes, that is, phrase types, phrase functions and part-of-speech tagging. The ETCBC database is freely available online and accessible with a Python package called text-fabric designed for searching and extracting data from this database (Roorda et al. 2018). Moreover, all programming codes and data for the present analysis is accessible online.¹

Accordingly, for the present analysis I have selected four complement types, the two of which are typically used as directional phrases and the remaining two as locative phrases. We assume that active verbs are more likely to combine with directional phrases and stative verbs with locative phrases.

Complement form	Hebrew	Presumed function
------------------------	---------------	--------------------------

¹ <https://github.com/ch-jensen/semantic-roles>

Preposition “in” + substantive	כּ	Locational
Prep. “on” + subs.	עַל	Locational
Prep “to” + subs.	אֶל	Directional
Subs. + morphological suffix	-הּ	Directional

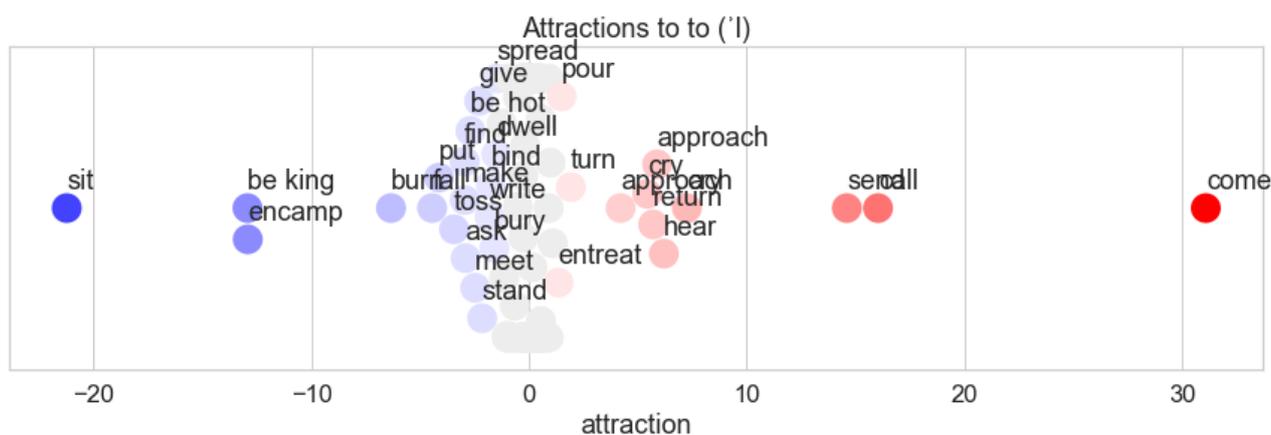
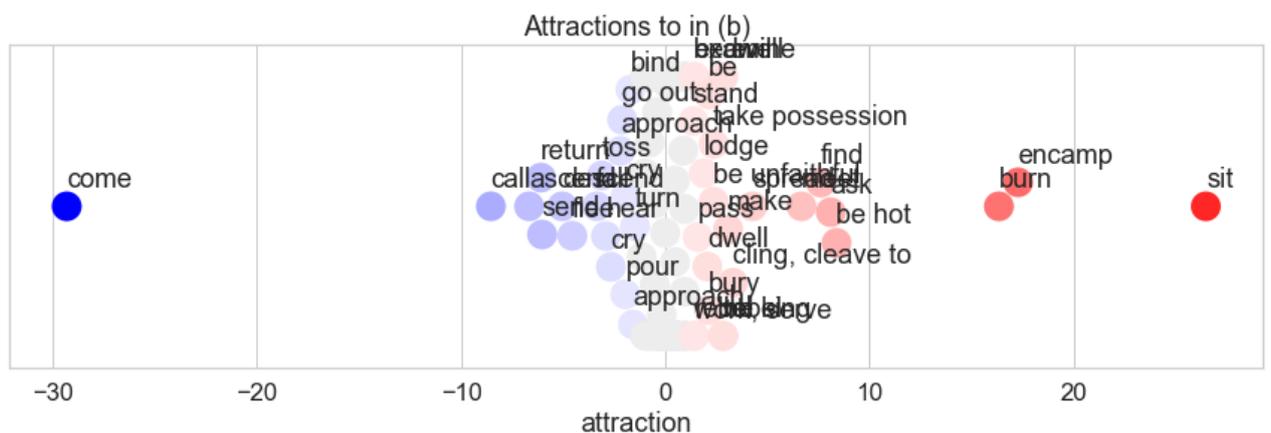
- (2) a. Abram lived in the land of Canaan (Genesis 13:12)
b. And they came to their father (Genesis 37:32)

It should be emphasized that prepositions are not semantically unambiguous. On the contrary, prepositions may often have different meanings depending on the verb. That is a paradox, since we want to use the complements as indicators of the aspectual meaning of the verb. To solve this issue, we need to classify or, at least, filter the complements. As an example, for the Hebrew preposition *B* which can both be used as a locative (“in”) and in instrument (“with”), I excluded all head nouns of the prepositional phrase which were likely to be used as instruments, such as “rock”, “stick”, and body parts. Other words are less likely to be instruments, such as human beings and geographical locations. In short, although the meaning of complement phrases is predicated by the verb, the internal composition of the complement may suggest a preferred use.

2.2 Results

The objective of the analysis is to compute the collocation strength between any given verb and any given adverbial, that is, how strongly the two components are associated in the corpus.² The plots show interesting associations. First, the stative verbs “sit” and “encamp” are very strongly associated with the preposition *B* “in”, while active verbs, most significantly “come”, are very weakly associated with this preposition. The other association plot, the plot of the preposition >L “to”, shows an almost opposite tendency. Here, “come” is very strongly associated with the preposition, while “sit” is weakly associated with the preposition. The two association plots support the notion of the complements being preferred for particular types of verbs. What is also important to note is the fact that most verbs are not particularly associated with either of these two complements, and the question remains to what extent will this type of analysis account for verbs in Biblical Hebrew.

² The scripts and layouts for the association plots here and the PCA below are adapted from Cody Kingham: https://github.com/CambridgeSemiticsLab/BH_time_collocations

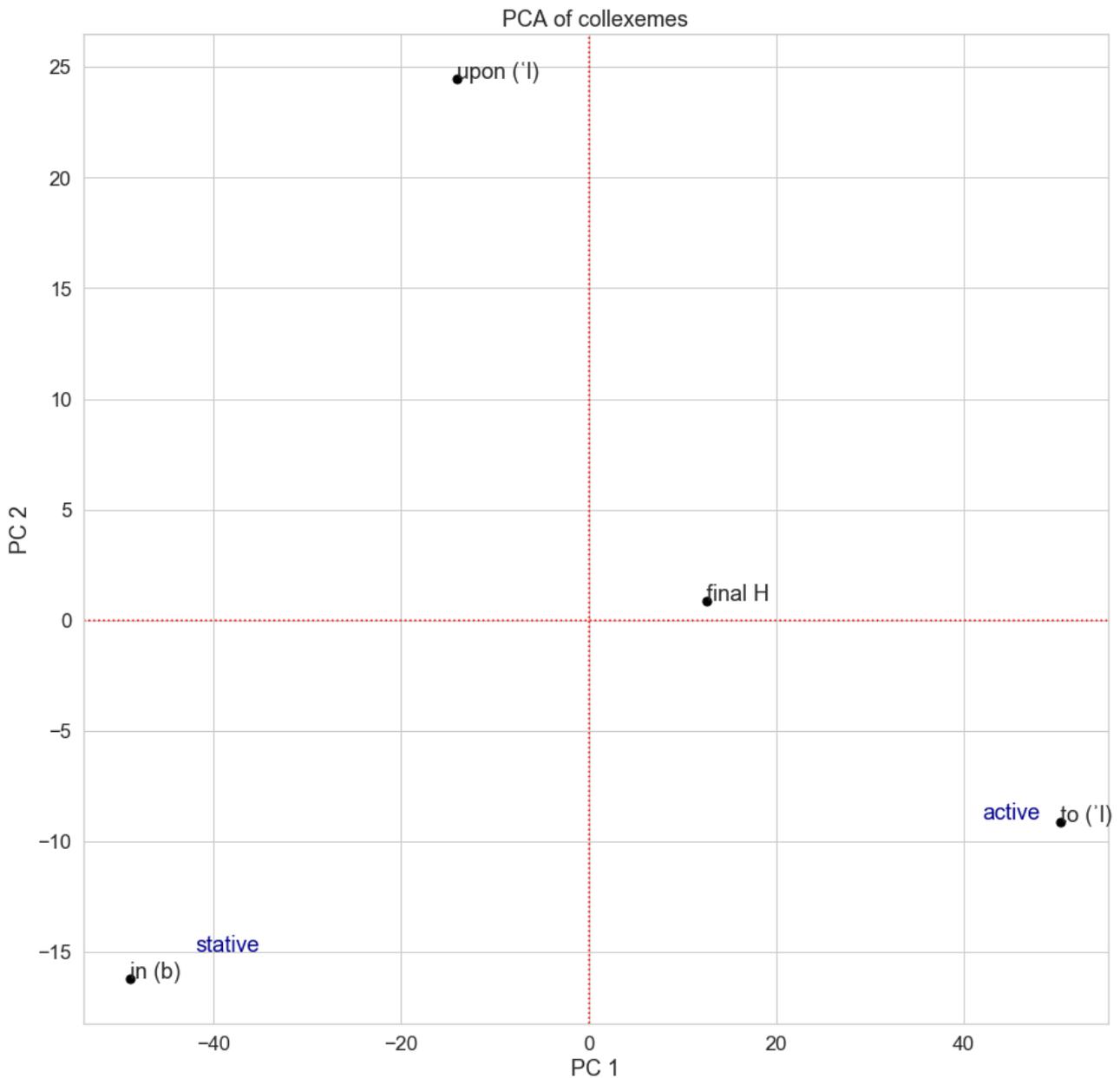


2.2.1 Principal Component Analysis

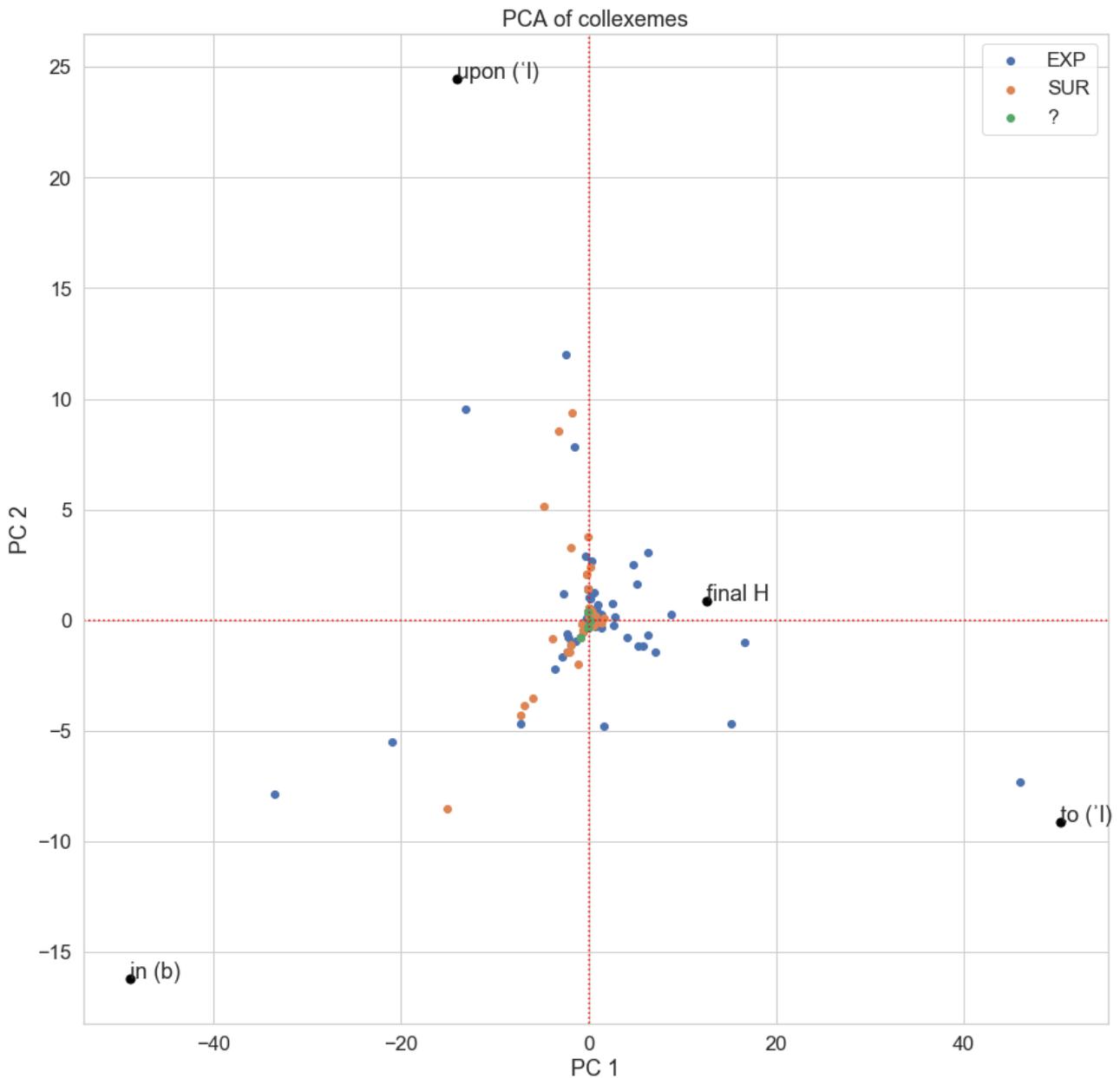
To get a clearer picture of how various verbs associate with various complement types, we can use a statistical method called Principal Component Analysis (PCA). PCA was developed to explore multiple independent quantitative variables and reduce the variation of these variables to the smallest possible number of dimensions, called “principal components”. The method has been widely used for a great diversity of types of data, including linguistic data.

The method also seems promising for the analysis of collocations of Hebrew verbs and their adverbials. A PCA was conducted of the four variables explored in this analysis and, as a matter of fact, 93.12% of the variation can be explained by the two first principal components. In other words, the large majority of the variation produced by four independent variables can be reduced to a two-dimensional plot.

The first principal component (PC 1) on the horizontal axis shows the largest variation formed by the two prepositions “in (b)” and “to (‘l)”. As a matter of fact, all four complement forms seem to support a variation of fientivity, the locative complements being on the left side of the plot and the directional forms on the right.



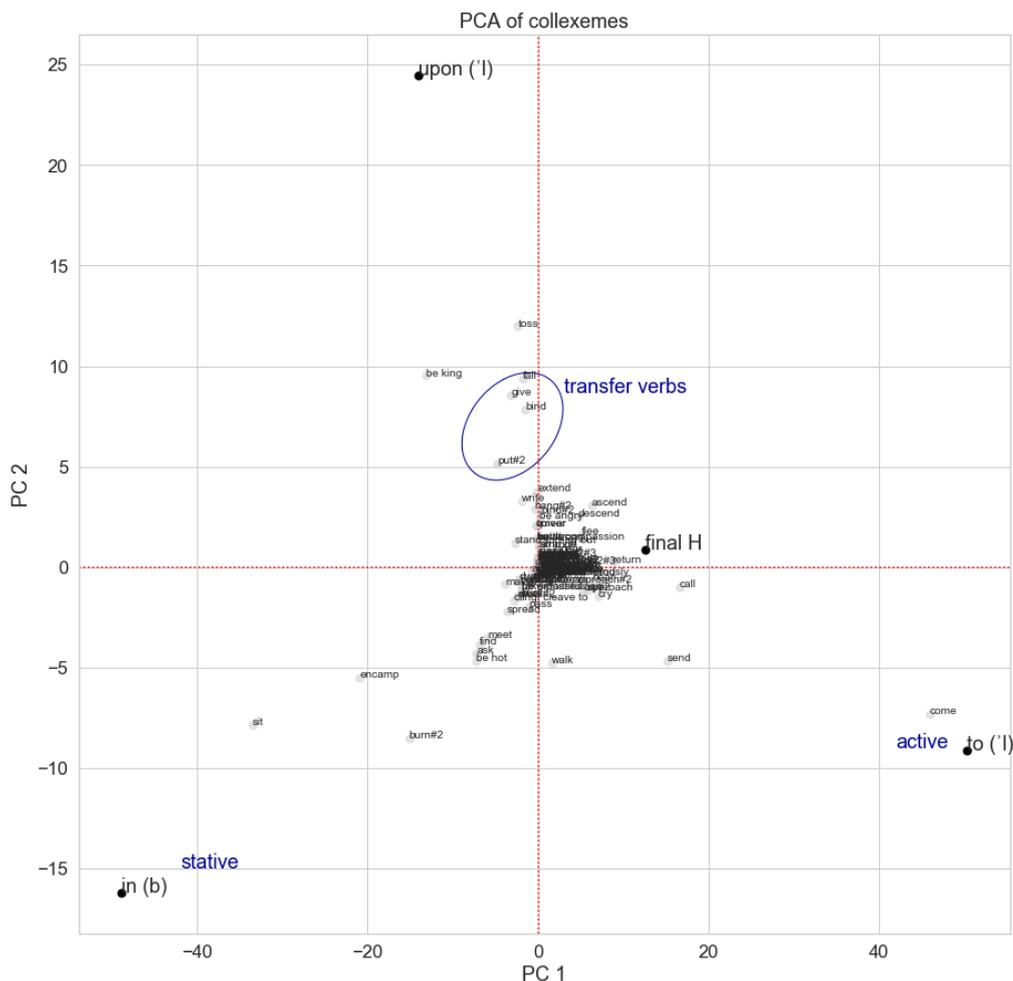
When plotting the individual lexemes onto the graph, this initial observation is confirmed: stative verbs, such as “sit”, “encamp”, and “be king”, dominate the left side of the plot while active (e.g. “come”, “send”, and “call”) verbs dominate the right side. Interestingly, “come” lies very closely to the complement form “to (‘l)” in contrast to all other motion verbs which are located near the centre of the plot. This distinctiveness of “come” matches an earlier study of the *Aktionsart* of Hebrew verbs conducted by Nicolai Winther-Nielsen who interpreted “come” as an Active Accomplishment. The close association between the verb and this complement type demonstrates that “come” occur in conjunction with a spatial endpoint of the movement, thus rightly seen as an Active Accomplishment.



A group of verbs seems to disturb the picture. These verbs include “give”, “put”, “take”, and “buy”. In one sense, they are all active in that they depict a transaction taking place over a span of time between two parties. On the other hand, these “transfer verbs” are more accurately seen as causative verbs of possession or location, depending on the type of transfer. A prototypical example is provided in Van Valin (2005, 61):

- (3) [do' (Pat, Ø) CAUSE [BECOME have' (Chris, book)]
 Pat gave the book to Chris

The lexical representation shows that “give” encompasses two events: 1) an active, inducing event, and 2) a stative, resultative situation. We may therefore expect the verb to occur in sentences with both stative and dynamic modifiers because each of the two subevents can be modified.



Apart from the dynamicity opposition, we also encounter an opposition between inherently causative verbs and non-causative verbs, and the PCA demonstrates our need for additional analytic tools. Therefore, I will now explore how we can use corpus linguistics to inform an RRG analysis of causative verbs.

3 Broadening the scope of analysis: Causatives

The canonical approach to testing verbs for causation has been to apply a test of paraphrasing. If a sentence can successfully be paraphrased into “cause + verb” without passivizing the verb or adding an additional participant, then the verb can safely be regarded as inherently causative (Van Valin 2005, 38). For Biblical Hebrew as well as other ancient languages, things are more complicated. If

we paraphrase Biblical Hebrew verbs, are we actually analysing the source language or our own translation of that language?

There is no simple syntactic test for testing causation. In fact, there may be no syntactic test at all to test for inherent causation in Biblical Hebrew. However, causativity correlates with transitivity in the sense that a causative is necessarily also transitive in that it involves two participants: a causer and a causee. So, I will now explore how a transitivity analysis may inform the decomposition of verbs as regards causativity.

3.1 Semantic transitivity

What is meant by “transitivity” here is not merely the number of participants. While the number of participants is certainly a prerequisite for talking about transitivity at all, transitivity is generally thought of as an activity that is “transferred” or “carried-over” from an agent to a patient (Hopper and Thompson 1980, 251). This “transfer” may be more or less effective depending on a variety of semantic parameters, as claimed by Hopper and Thompson (1980) in transitivity Hypothesis. Among the semantic features important for grasping the degree of transitivity are dynamicity (called “kinesis” by Hopper and Thompson), telicity (“aspect”) and volitionality, that is, whether the Agent acts by purpose or by accident.

With respect to causation, Hopper and Thompson posit the following definition seen within the framework of their Transitivity Hypothesis (1980, 264):

“... causatives are highly Transitive constructions: they must involve at least two participants, one of which is an initiator, and the other of which is totally affected and highly individuated.”

With this definition in mind, it is worth exploring whether the degree of affectedness of the participants may help determining if a verb is causative or not. More recently, Hopper and Thompson’s Transitivity Hypothesis has been re-addressed in Åshild Næss’ *Prototypical Transitivity*, and it is her framework that I have used for my research.

3.2 Affectedness

One of the advantages of Næss’ definition of semantic transitivity is its simplicity. Næss defines the prototypical transitivity as a transitive construction where the Actor and the Undergoer are “maximally distinguished” (Næss 2007). The greatest distinction arises where the Actor controls the event, or intentionally instigates the event, in Næss’ terms, and where the Undergoer is totally affected while not being neither volitional nor instigation (cf. Næss 2007, 44):

	Agent	Patient
Volitionality	+	–
Instigation	+	–
Affectedness	–	+

Næss herself does not discuss the relationship between transitive and causative constructions, but it seems to me that this definition of prototypical transitivity corresponds to causation. In particular, it seems that the Affectedness opposition is crucial for defining causation. If the Actor is not affected by the event, we can assume the event to be causative. And I will now test this assumption on a number of propositions from the corpus.

Although Næss’ approach differs from an RRG approach, she reaches a number of interesting conclusions that intersects with canonical RRG. The first of these regards verbs of Emotions:

- (4) ʔiš ʔim -mô w^o ʔāv -îw tîrā -ʔû
man mother -PRS.M.3SG CON father-PRS.M.3SG fear.IMPF -M.2PL

“Everyone shall fear his father and mother” (Leviticus 19:3)
fear’ (everyone, his father and mother)

The sentence in (4) is minimally transitive because the syntactic object is not affected by the event. By contrast, the subject is volitional and affected, thereby labelled “Volitional Undergoer”. In RRG the verb would be seen as a stative, non-causative verb.

- (5) šivʕat yām -îm mas -šôt tōʕxēl -û
seven day -M.PL bread -F.PL eat.IMPF -M.2PL

“For seven days you shall eat unleavened bread” (Leviticus 23:6)
do’ (you, [**eat’** (bread)])

One might be tempted to treat (5) as a causative construction because the Undergoer is clearly affected. However, in RRG consumption verbs are treated as simple activities, presumably because it would be illegitimate to paraphrase the sentence as “you shall cause the unleavened bread to be eaten” which would passivize the verb. Næss’ analysis adds an additional reason: The sentence is not causative, because the Actor is itself affected by eating it. One of the main criteria for defining affectedness is change and (temporal or geographical) scope. In this case, the temporal scope would normally be constrained by the Actor who eats until he or she has become full. In this respect, the sentence is less transitive and therefore unlikely to be causative.

Also construction verbs are seen as less transitive, and they would similarly be seen as non-causative in correspondence with RRG. In contrast to consumption verbs, however, the Actor is not affected, so the constraint lies with the Undergoer. In this case because the Undergoer is not undergoing a change as such, but as a matter of fact comes into existence, it would be more accurate to see the Undergoer as being *effected* rather than *affected*, as in (6):

(6) û ntaʕ -tem kol ʕēš maʔaxāl
 CON plant.PERF-M.2PL any tree eat

“and you plant any eatable tree” (Leviticus 19:23)
do’ (you, [**plant’** (any eatable tree)])

As we have seen, the notion of Affectedness allows us to identify non-causative verbs. We can also use the criterion positively to identify causative verbs, such as in (7):

(7) w° lōʔ t° -hallēl ʔet šēm ʔlōheʔ -xā
 CON NEG M.2SG -defile.IMPF OBJ name God -PRS.M.2SG

“so that you defile the name of your God” (Leviticus 18:21)
[Do’ (you, Ø)] CAUSE [BECOME **unclean’** (the name of your God)]

In (7) the Actor is not affected by the event, unlike the Undergoer which is heavily affected in that it becomes impure. The notion of Affectedness is also useful when decomposing a transfer verb like “give”:

(8) hā ʔāreš ʔašer ʔanî nōtēn lā -xem
 DET land REL I give.PTC to -PRS.M.2PL
 “the land which I have given to you” (Leviticus 25:2)
[Do’ (I, Ø)] CAUSE [BECOME **have’** (you, the land)]

The ditransitive clause in 0 has three participants, the subject (“I”) the object (“the land”) and the dative (“you”), all of which are affected by the transfer. Obviously, one may discuss whether “the land” can really be affected. On the other hand, “the land” is regularly portrayed as the possession of God (the Actor) who can give and take away the land. In any case, the example shows that it is not always simple to encapsulate Affectedness. Even though Næss presents it as a binary category, many factors play into this concept, and Næss also admits that Affectedness is rather a continuum.

4 Conclusion

In conclusion, an RRG approach to lexical decomposition benefits from incorporating corpus linguistics and related methods, in particular when analysing the language of ancient corpora where

the researcher cannot rely on his own intuition for the language. In this paper, I have demonstrated how covarying collexeme analysis can be successfully applied to Biblical Hebrew in order to determine the fientivity of Biblical Hebrew verbs on a quantitative basis. Secondly, I explored how a semantic tagging of Affectedness may help distinguishing causative from non-causative transitives. Obviously, much research still needs to be done, especially as regards a development of more fine-grained annotations of complement types as well as of participant roles. I believe, however, that the research carried out so far demonstrates the fruitfulness of applying quantitative methods to inform a Role and Reference Grammar of Biblical Hebrew.

Bibliography

- Dowty, David R. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Synthese Language Library 7. Dordrecht ; Boston: D. Reidel Pub. Co.
- Gries, Stefan Th., and Anatol Stefanowitsch. 2004. "Covarying Collexemes in the Into-Causative." In *Language, Culture, and Mind*, edited by M. Achard and S. Kemmer, 225–36. Stanford, CA: CSLI.
- Hopper, Paul J., and Sandra A. Thompson. 1980. "Transitivity in Grammar and Discourse." *Language* 56 (2): 251–99.
- Næss, Åshild. 2007. *Prototypical Transitivity*. Typological Studies in Language 72. Amsterdam: John Benjamins Publishing.
- Roorda, Dirk, Cody Kingham, Christiaan Erwich, and SeHoon Park. 2018. "Etcbc/Bhsa: Default Text Format For Lexemes." Zenodo. <https://doi.org/10.5281/zenodo.1446393>.
- Stefanowitsch, Anatol, and Stefan Th. Gries. 2005. "Covarying Collexemes." *Corpus Linguistics and Linguistic Theory* 1: 1–43.
- Van Valin, Robert D., Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.
- Winther-Nielsen, Nicolai. 2016. "How to Classify Hebrew Verbs: Plotting Verb-Specific Roles." In *Contemporary Examinations of Classical Languages (Hebrew, Aramaic, Syriac, and Greek): Valency, Lexicography, Grammar, and Manuscripts*, edited by Timothy Martin Lewis, Alison Salvesen, and Beryl Turner, 67–94. Perspectives on Linguistics and Ancient Languages 8. Piscataway, NJ: Gorgias Press.