MACULA Hebrew Treebank for Open Scriptures Hebrew Bible (OSHB) Initial Release Documentation

Principal Authors: Randall Tan and Andi Wu

March 31, 2022

1 Introduction

Since the early 2000s, Clear Bible, Inc. (formerly known as Global Bible Initiative [2014-2020] and Asia Bible Society [before 2014]) has been richly annotating the original Greek and Hebrew texts of Scripture. This document describes the MACULA Hebrew Treebank for the Open Scriptures Hebrew Bible (OSHB), which is released under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

While the MACULA OSHB Treebank is a new project, its various components have a long history. The text and morphology come from the Open Scriptures Hebrew Bible (OSHB) project, which began in 2009 and is licensed CC BY 4.0. The text used in OSHB is the Westminster Leningrad Codex (WLC), which was first completed in 1987 and is in the public domain. In addition, Clear has previously collaborated with The J. Alan Groves Center for Advanced Biblical Research ("The Groves Center" for short) on the Westminster Hebrew Syntax (WHS). The WHS was generated initially using Clear's automatic parser and machine-readable Hebrew grammar together with the Westminster Hebrew Morphology (WHM) on the Westminster Leningrad Codex (WLC). Clear then took primary responsibility for manually checking and editing the trees. Recently, the Groves Center graciously released Westminster Hebrew Syntax without Morphology under CC BY 4.0.

The MACULA OSHB Treebank uses the same tree structure as the WHS, with OSHB morphology replacing the Westminster morphology. Because the trees initially analyzed the Westminster Leningrad Codex (WLC) and Westminster Hebrew Morphology (WHM), systematic adjustments have been made (and still need to be made) to the trees to reflect the use of the OSHB text and morphology instead. For this initial release, a number of discrepancies in analysis between the OSHB morphology and the syntactic representation have not yet been fully manually checked. We expect revisions for both specific interpretational differences and systematic differences in representation will still need to be made. Aside from new issues stemming from switching over to the OSHB morphology, continued work is needed for the Hebrew Treebank itself to remove any remaining inconsistencies and errors. Where multiple interpretations are possible, alternative annotations still need to be added. We welcome user feedback and alternative analyses to improve it.

2 Overview

The treebank is designed to be eclectic and flexible to maximize potential compatibility with different linguistic frameworks. We purposely use structures that form the common grounds of linguistic theories rather than strictly adhere to any one approach. We also strived to use the same or similar syntactic representations for our Greek and Hebrew treebanks to allow for easier comparative analysis across the whole Bible. However, significant differences remain. For example, the Greek trees are sentence-based and frequently cross verse boundaries, whereas the Hebrew trees are verse-based.

2.1 Nodes and Attributes

Each node in a tree has zero or more child nodes. Terminal nodes are nodes that have no child nodes. All nodes in a tree that have child nodes are non-terminal nodes. The following table describes attributes in all terminal and/or non-terminal nodes of the trees. Besides the trees, a preliminary form of some MACULA datasets are being made available separately among the sources published together with the

trees. We have included explanations of these annotations and what they mean at the bottom of the table. They are marked with "(soon to be integrated)." These will soon be integrated into the trees, but did not make it into this initial release. For those would benefit from this data sooner, we wanted to make them available now.

Attribute	Applies To	
ID	Each Sentence	Should be treated as a string. Normally corresponds to the book, chapter, and verse number.
nodeId	Terminals Non-terminals	A 15-digit unique id in the format BBCCCVVVWWWSSSL where BB => zero-padded book CCC => zero-padded chapter VVV => zero-padded verse WWW represents the beginning position (the Nth word) of a node/sub-tree SSS represents the SPAN of a node (how many terminal nodes it covers) L (Level) is used to distinguish nodes which have the same span (in cases of non-branching nodes)
Cat	Terminals Non-terminals	Syntactic category. See sections 2.2, 2.3, and 2.4 below for details.
Rule	Non-terminals	Label of the rule used to derive the non-terminal node.
Head	Non-terminals	The 0-based index indicating the position of the "Head" node among its child nodes.
n (identical to morphId in earlier versions)	Terminals	An 12-digit unique id in the format BBCCCVVVWWWP where BB => zero-padded book CCC => zero-padded chapter VVV => zero-padded verse WWW => zero-padded word index (for the word instance within the verse) P -> word part (when a word written together is comprised of multiple parts that play different syntactic functions) Note: When phrase or clauses have intervening elements outside of the phrase or clause, words have to be moved in word order to enable the phrase or clause to be represented in the typical way. n keeps track of the original word order even in such circumstances.
m elements	Terminals	Elements that contain the morphemes of the Hebrew text.
c elements	Terminals	Elements that contain multiple morphemes that make up a compound word. OSHB sometimes analyzed the constituent parts of compound words in a way that caused a conflict with the syntactic function. c elements group such morphemes into one terminal node in the tree.
lang	Terminals (within m)	Used to distinguish Hebrew and Aramaic words. The vast majority of words have the value "H" to indicate "Hebrew," while words in Aramaic are "A."
morph	Terminals (within m)	Copy of the code OSHB used for parsing the Hebrew text. See https://hb.openscriptures.org/parsing/HebrewMorphologyCodes.html.
pos	Terminals (within m)	Part of speech according to OSHB. See https://hb.openscriptures.org/parsing/HebrewMorphologyCodes.html.
type	Terminals (within m)	Distinguishes different types of verb conjugations, adjectives, nouns, pronouns, suffixes, and particles. Follows OSHB. See https://hb.openscriptures.org/parsing/HebrewMorphologyCodes.html.
stem	Terminals (within m)	Indicates both the kind of verbal action (simple, stative, causative) and the voice (active, passive, reflexive) of the verb. Follows OSHB..
person	Terminals (within m)	Indicates the participant of an event as first, second, or third person for pronouns, pronominal suffixes, and verbs. Follows OSHB. See https://hb.openscriptures.org/parsing/HebrewMorphologyCodes.html.

gender	Terminals (within m)	Grammatical gender, which forms an agreement system with adjectives, nouns, pronouns, pronominal suffixes and verbs. Follows OSHB. See https://hb.openscriptures.org/parsing/HebrewMorphologyCodes.html.
number	Terminals (within m)	Grammatical number, which forms an agreement system with adjectives, nouns, pronouns, pronominal suffixes and verbs. For example, count nouns have a singular form referring to one object and a plural form referring to multiple objects. Some nouns also have a dual form. Follows OSHB. See https://hb.openscriptures.org/parsing/HebrewMorphologyCodes.html.
state	Terminals (within m)	Nouns, adjectives, participles and infinitives can appear in either the absolute state or the construct state. The absolute state is the standard form. A word that takes a suffix or is connected to another term in a construct chain is in the construct state. Follows OSHB. See https://hb.openscriptures.org/parsing/HebrewMorphologyCodes.html.
Greek (soon to be integrated)	Terminals	Transliteration of the corresponding Greek for the Hebrew word through a tentative alignment to the Septuagint. Blank when the Hebrew word did not align to any Greek word in the Septuagint for one of the following reasons: (1) the Hebrew word didn't get translated in the Septuagint; or (2) it failed to be aligned. This data has never been manually checked and is released as a starting point for further work.
GreekStrong (soon to be integrated)	Terminals	Strong Number of the corresponding Greek for the Hebrew word through a tentative alignment to the Septuagint. Value is 0 when the Hebrew word did not align to any Greek word in the Septuagint for one of the following reasons: (1) there is no corresponding word in the NT and therefore there is not a StrongNumber for that; (2) the Hebrew word didn't get translated in the Septuagint; or (3) it failed to be aligned. This data has never been manually checked and is released as a starting point for further work.
SenseNumber (soon to be integrated)	Terminals	Tentative probabilistic grouping of the different senses of the Hebrew word: (1) the grouping is created by a data-driven clustering algorithm where the granularity can be adjusted; (2) It is different from a sense dictionary in that it is token-based rather than type-based, i.e. it doesn't just provide the sense groups of a word, but assigns a particular sense to each instance of the word; (3) it can be used in conjunction of a sense dictionary to produce an assignment of senses for each occurrence of the Hebrew words. Value is 0 for function words and certain other high frequency words like the "to be" verb, where no attempt was made to group into different senses. Words with three senses would have SenseNumber 1, 2, and 3 to distinguish the three senses in its various occurrences. Words with just one sense would only have SenseNumber 1 for all occurrences. This data has never been manually checked and is released as a starting point for further work.
Ref (soon to be integrated)	Terminals	Manually-checked dataset that disambiguates the referents of pronouns. Contains values only for pronouns. The value represents the morphId (inside curly brackets {}) of the nearest nominal antecedent of a pronoun in context. For example, in Genesis 1:27 "God" (with morphId 010010270021) is the nearest nominal antecedent of the pronoun "his" at morphId 010010270053. So, the value of {010010270021} is in the Ref attribute for the pronoun "his" at morphId 010010270053 to indicate that "God" is the pronoun's nominal antecedent.
SubjRef (soon to be integrated)	Terminals	Manually-checked dataset that disambiguates the implied subjects of verbs. Contains values only for certain verbs with an implied subject. The value represents the morphId (inside curly brackets {}) of the implied subject in context. For example, in Genesis 1:5 "God" (with morphId 010010050021) is the implied subject in context of the verb "he called" (at morphId 010010050061). So, the value of {010010050021} is in the SubjRef attribute for the verb "he called" at morphId 010010050061 to indicate that "God" is the verb's implied subject.

Frame (soon to be integrated)	Terminals	Partially checked, tentative dataset that adds semantic information on predicate arguments as inspired by the original PropBank project (http://propbank.github.io). In the Frame attribute of verbs, the morphId of their agents and patients are entered. A0 indicates arguments exhibiting features of prototypical Agents. A1 indicates arguments exhibiting features of prototypical Patients. AA indicates Agents of induced actions—cause A0 to act. For example, in Genesis 1:1, the verb "he created" (with morphId 010010010021) has the value of {A0:010010010031; A1:010010010052;010010010072;} to indicate "God" (morphId 010010010031) as the Agent and the "heavens" (morphId 010010010052) and the "earth" (morphId 010010010072) as the Patient.

2.2 Syntactic Categories at Word Level: Part of Speech Labels

A terminal node is the basic unit of syntactic analysis, usually corresponding to the type of speech analysis for each word as provided by the OSHB morphology. OSHB distinguishes nine categories of part of speech (https://hb.openscriptures.org/parsing/HebrewMorphologyCodes.html). The following table spells out the labels used for these nine parts of speech in the treebank:

Label	Part of Speech Spelled Out
adj	adjective
adv	adverb
conj	conjunction
noun	noun
prep	preposition
ptcl	particle
pron	pronoun
sfx	suffix
verb	verb

The treebank also currently distinguishes five additional categories:

Label	Part of Speech Spelled Out	Mapping to OSHB
om	object marker	particle type direct object marker
art	definite article	particle type definite article
ij	interjection	particle type interjection, as well as exhortation, demonstrative
rel	relative particle	particle type relative
num	numeral	adjective type cardinal number, ordinal number

OSHB has five particle types (affirmation, exhortation, interrogative, demonstrative, and negative) that do not currently correspond to an existing terminal node part of speech. Interjection is also frequently used to cover what OSHB considers exhortation and demonstrative particles. These discrepancies require more careful analysis and will be revisited in the future.

Adjective (adj): A word that belongs to a class whose members modify nouns. An adjective specifies the properties or attributes of a noun it modifies.

Adverb (adv): Narrowly defined, a word that belongs to a class of words whose members modify verbs for such categories as time, place, manner, direction, etc. Broadly defined, includes words that modify any constituent class of words, such as verbs, adjectives, adverbs, phrases, clauses, or sentences. Some examples of this broader class are degree words and negatives.

Conjunction (conj): A word that syntactically links words or larger constituents and expresses a semantic relationship between them.

Noun (noun): A word that belongs to a class whose members most frequently act as the subjects, objects or indirect objects of the verb or the object of a preposition. Includes words that typically refer to concrete or abstract entities like people, places, things and concepts.

Preposition (prep): A word that occurs before a nominal phrase, forming a single unit with it (a prepositional phrase) to express the prepositional phrase's semantic relation to another unit within the clause (typically indicating when, where, how or why).

Particle (ptcl): A catchall term for a variety of function words. A number of OSHB particle types currently are represented generally as particles while others (e.g., object marker, definite article, relative particle) have been specifically named. Particles are typically invariable in form and some partially overlap with the broader definition of adverbs that modify at the clause or sentence level.

Pronoun (pron): A word substituting for a noun or nominal phrase whose referent is recoverable from the linguistic or extralinguistic context.

Suffix (sfx): Biblical Hebrew uses four different types of suffixes, which are letters appended to the end of a word. The most common type is the pronominal suffix.

Verb (verb): A word that belongs to a class whose members typically indicate events or actions. It governs the number and types of other constituents that may occur in a clause.

Object marker (om): A particle immediately precedes the noun or phrase that functions as the direct object of a verb.

Definite article (art): A particle that marks a noun or phrase as definite.

Interjection (intj): A particle that expresses strong emotion. Subject to future reevaluation, for the moment particles that strengthen the emotion of a request or command (exhortation particles) and particles that direct a listener or hearer's attention (demonstrative particles) have been lumped with interjection particles in the current trees.

Relative particle (rel): A particle that introduces relative clauses.

Numeral (num): Cardinal numbers and ordinal numbers are two types of attributive adjectives that can be classed as numerals. Cardinal numbers express the name of a number itself or the quantity of something. Ordinal numbers express the rank or order of things in a series (first, second, third, etc.) or a part of a whole (a third, a fourth, a fifth, etc.).

2.3 Syntactic Categories at Phrase Level

The phrase level is the intermediate level between word level and clause level. Phrase level nodes are either non-terminal nodes that are the immediate parent nodes of the part-of-speech terminal nodes or parent nodes of other phrase level non-terminal nodes that together form multi-word phrases. From the perspective of the clause, single words or combinations of words form phrases, which are the minimal constituents with a specific function at the clause level. Only seven categories of phrases are distinguished, using the following labels: adjp, advp, ijp, np, pp, relp, vp. (Note: cjp and omp are slated for imminent removal and have no meaningful content—simply use the cj and om word level meanings)

The following table spells out the labels used for the seven categories of phrases distinguished:

Label	Phrase Category Spelled Out
adjp	adjectival phrase
advp	adverbial phrase
ijp	interjectional phrase
np	nominal phrase
pp	prepositional phrase
relp	relative phrase
vp	verbal phrase

Adjectival Phrase (adjp): A phrase with an adjective as its head. Adjectival phrases typically function as dependent modifiers within a nominal phrase. In many cases, an adjectival phrase functions in place of an elided noun and thus functions as a nominal phrase (this is the reason why adjectival phrases are often promoted to nominal phrases in the tree).

Adverbial Phrase (advp): A phrase with an adverb as its head. Adverbial phrases typically have Adverbial function at the clause level, indicating the "when", "where", "why" and "how" of the verb or non-verbal predicate. Some adverbial phrases also modify nouns or adjectives.

Interjectional Phrase (ijp): Currently conflates phrases that has either a particle that expresses strong emotion or a particle that draws attention to what immediately follows it. (Likely needs to be distinguished in future releases.)

Nominal Phrase (np): A phrase with a noun (or adjective or pronoun functioning as a noun) as its head. Nominal phrases typically have Subject or Object function at the clause level. Nominal phrases are also often modified by a preposition, forming a prepositional phrase.

Relative Phrase (relp): A phrase introduced by a relative particle (really a relative clause).

Prepositional Phrase (pp): A phrase with a preposition and the nominal phrase it governs, forming a single unit with the nominal phrase to express semantic relation to another unit within the clause (typically indicating when, where, how or why). Prepositional phrases typically have an Adverbial function at the clause level. Some prepositional phrases modify a noun or adjective instead of a verb (indicating a relationship of where, when, how, why for the object of the preposition to the head noun or adjective).

Verbal Phrase (vp): A phrase with a verb as its head.

2.4 Syntactic Categories at Clause Level

The clause level differs from the phrase level by using a dependency-like structure. The terminology used to describe the functions of clause level constituents is purposely conservative for ease of understanding and to preserve a clearer link between clause level terminology and phrase and word level terminology. In general, different parts of speech have the following clause-level functions: Verbs have a Verbal Function. Unlike Greek, verbal copulas are not used in Hebrew, which rely on verbless clauses exclusively for "to be" clauses. Nominals (i.e., nouns and other parts of speech that can function as a noun) can have four basic functions in relation to the verb: Subject, Object, or Adverbial. Nominals can also function as a Predicate to either the Subject (i.e., as Predicate in Verbless Clauses) or the Object (i.e., as Object Complement, considered a Second Object). Adverbs and prepositional phrases at the clause constituent level function adverbially in relation to the verb. Eight categories of clause level function are distinguished, using the following labels: ADV, O, O2, OC, P, S, V, PP.

The following table spells out the labels used for the eight categories of clause functions distinguished:

Label	Part of Speech Spelled Out
ADV	Adverbial Function
O	Object Function
O2	Second Object Function
OC	Object Complement Function
S	Subject Function
P	Predicate Function
V	Verbal Function
PP	Prepositional Phrase Function

Adverbial Function (ADV): A constituent that represents when, where, how, or why of a proposition.

Object Function (O): A constituent that represents the patient or goal of the action of a proposition.

Second Object Function (O2): Some verbs take two objects. There are two main types. The first type involves an object of person (the first object) and an object of thing (the second object).The second type involves a direct object and an object complement. The object complement predicates a description of the direct object (e.g., "king" is the object complement in "God appointed David as king"). In the current version of the trees, these two types are not yet distinguished.

Object Complement Function (OC): A constituent that makes an assertion about the object. (The distinguishing of object complements (OC) is planned for the near future.)

Subject Function (S): A constituent that represents the agent of typically transitive verbs and the single argument of intransitive verbs.

Predicate Function (P): A constituent that makes an attribution or identification about the Subject of a verbless clause or a verbal clause with a verbal copula.

Verbal Function (V): A constituent that represents the action/event of a proposition. In verbal clauses, this is the head of the clause, on which all other clause constituents depend.

Prepositional Phrase Function (PP): A modifying phrase consisting of a preposition and its object that represents when, where, how, or why of a proposition.

3 Annotation Style

In this documentation, where applicable, a conscious attempt was made to address the same general areas in sequence as the MACULA Nestle 1904 Greek New Testament Treebank Guidelines. Examples are taken from Genesis exclusively and are usually short and focused on the construction at hand.

3.1 Verbs (V) and Non-Verbal Predicates (P)

Every verb, finite or non-finite, is automatically promoted to a vp at the phrase level and V at the clause constituent level, except for Infinitive Absolute verbs, which are promoted to an advp and ADV respectively. Verbs can also form clauses on their own. Verbs are not labeled as predicates or predicators in the trees, but are seen as the head/core constituent of each clause of which it is a part.

Example 1 (from Genesis 1:28): a single finite verb (V) as vp, V, and then CL:

[CL]
[V]

[vp]
[verb] פְּרוּ (be fruitful!)

Example 2 (from Genesis 37:16): a single finite verb (V) as vp, V, and then CL:

[CL]
[O]
[np]
[omp]
[om] אֶת־ (-)
[np]
[np]
[noun] אַח (brothers of)
[np]
[pron]ִי (me)
[S]
[np]
[pron] אָנֹכִי (I)
[V]
[vp]
[verb] מְבַקֵּשׁ (looking)

Example 3 (from Genesis 2:17): Infinitive Absolute verb as ADV and finite verb as V in an PP-ADV-V clause, plus Infinitive Construct verb as V in an V-S-PP embedded clause that modifies a noun:

[CL]
[PP]
[pp]
[pp]
[prep] בְּ (on)
[np]
[np]
[noun] יוֹם (day)
[CL]
[V]
[vp]
[verb] אֲכָל (to eat)
[S]
[np]
[pron] ךָ (you)
[PP]
[pp]
[pp]
[prep] מִמּ (from)
[np]
[pron] ֶּנּוּ (it)
[ADV]
[advp]
[verb] מוֹת (to die [affirmation function: surely])
[V]
[vp]
[verb] תָּמוּת (you will die)

Example 4 (from Genesis 1:2): a prepositional phrase as a non-verbal predicate P in a verbless S-P clause:

[CL]
[S]
[np]
[noun] חֹ֫שֶׁךְ (darkness)
[P]
[pp]
[prep] עַל (over)
[np]
[np]
[noun] פְּנֵי (surface)
[np]
[noun] תְּהֹ֑ום (of deep)

Every complete clause (i.e., non-elliptical) has a verb (V) or a non-verbal predicate (P). There are also clauses without any predication—no verb (V) or non-verbal predicate (P). This yields 3 types of clauses:

1. Complete clauses with a verb (V) are assumed to be Verbal Clauses (e.g., examples 1, 2, and 3 above).

2. Clauses with the verb elided have the Type attribute Verb Elided. Example 5 (from Genesis 30:42): The second clause, a S-PP verb elided clause, is interpreted as assuming a form of the verb "to be" from the previous V-S-PP clause:

[cjp]
[cj] וְ (so)
[CL]
[V]
[vp]
[verb] הָיָה (was)
[S]
[np]
[art] הָ (the)
[np]
[noun] עֲטֻפִים (ones being weak)
[PP]
[pp]
[pp]
[prep] לְ (to)
[np]
[noun] לָבָן (Laban)
[cjp]
[cj] וְ (and)
[CL]
[S]
[np]
[art] הַ (the)
[np]
[vp]
[verb] קְשֻׁרִים (ones being strong)
[PP]
[pp]

[pp]
[prep] לְ (to)
[np]
[noun] יַעֲקֹב (Jacob)

3. Clauses with a non-verbal predicate (P) are Nominal Clauses (e.g., example 4).

4. Vocatives, left-dislocated focus noun phrases, and some interjection clauses are categorized as Nominal Clauses for now. In the Greek trees, these are Minor Clauses. For the sake of consistency, these cases may be changed to Minor Clause in the future. Example 6 (from Genesis 4:23): Vocative clause shown together with the main clause to which it is connected:

[CL]
[CL]
[np]
[np]
[noun] עָדָה (Adah)
[cjp]
[cj] וְ (and)
[np]
[noun] צִלָּה (Zillah)
[CL]
[V]
[vp]
[verb] שְׁמַעַן (listen!)
[O]
[np]
[np]
[noun] קוֹל (voice of)
[np]
[pron] יִ֫ (me)

Example 7 (Genesis 34:8): Left-dislocated noun phrase shown together with the main clause to which it is connected:

[CL]
[CL]
[np]
[np]
[noun] שְׁכֶם (Shechem)
[np]
[np]
[noun] בְּנ (son of)
[np]
[pron] יִ֫ (me)
[CL]
[V]
[vp]
[verb] חָשְׁקָה (is set)
[S]
[np]
[np]
[noun] נַפְשׁ (soul of)

[np]
[pron] וֹ (him)
[PP]
[pp]
[pp]
[prep] בְּ (on)
[np]
[np]
[noun] בַּת (daughter of)
[np]
[pron] כֶּם (you)

3.2 Subject (S)
Explicit subjects of verbs are labeled as S (whereas the subject of the majority of clauses is implied through the person and number of the verb and is not explicitly stated). Nouns, pronouns, and other parts of speech functioning as a noun typically serve as subjects of finite verbs.

3.2.1 Nouns
Example 8 (Genesis 1:3): The noun is the subject of the main verb:
[cjp]
[cj] וַ (and)
[CL]
[V]
[vp]
[verb] יֹּאמֶר (said)
[S]
[np]
[noun] אֱלֹהִים (God)

3.2.2 Pronouns
Example 9 (Genesis 3:19): The pronoun is the subject of the main verb:
[CL]
[P]
[np]
[noun] עָפָר (dust)
[S]
[np]
[pron] אַתָּה (you)

3.2.3 Infinitive clauses
Example 10 (Genesis 2:18): The infinitive clause as a whole is the subject of the main verb:
[CL]
[ADV]
[advp]
[adv] לֹא־ (not)
[P]
[adjp]
[adj] טֹוב (good)
[S]
[np]
[CL]

[V]
[vp]
[verb] הֱיוֹת (to be)
[S]
[np]
[art] הָ (the)
[np]
[noun] אָדָם (man)
[PP]
[pp]
[pp]
[prep] לְ ()
[np]
[np]
[noun] בַּד (alone)
[np]
[pron] וֹ ()

3.3 Object (O)
Like subjects, objects may not be explicitly stated (i.e., they may be elided). Nouns, pronouns, and other parts of speech functioning as a noun typically serve as objects of verbs.

3.3.1 Nouns
Example 11 (from Genesis 1:7): The noun is the object of the main verb:

[cjp]
[cj] וְ (and)
[CL]
[O]
[np]
[art] הַ (the)
[np]
[noun] דֶּלֶת (door)
[V]
[vp]
[verb] סָגַר (he shut)
[PP]
[pp]
[pp]
[prep] אַחֲרֵי (after)
[np]
[pron] ו: (him)

3.3.2 Object markers
Example 12 (from Genesis 1:7): The object marker frequently marks the definite direct object of transitive verbs:

[cjp]
[cj]
[cj] וַ (so)
[CL]
[V]
[vp]

[verb] יַּעַשׂ (he made)
[S]
[np]
[noun] אֱלֹהִים (God)
[O]
[np]
[omp]
[om] אֶת־ (-)
[np]
[art] הָ (the)
[np]
[noun] רָקִיעַ (expanse)

3.3.3 Infinitival clauses

Example 13 (from Genesis 4:12): The infinitival clause as a whole is the object of the main verb:

[CL]
[ADV]
[advp]
[adv] לֹא־ (not)
[V]
[vp]
[verb] תֹּסֵף (she will continue)
[O]
[CL]
[V]
[vp]
[verb] תֵּת־ (to yield)
[O]
[np]
[np]
[noun] כֹּחַ (crop of)
[np]
[pron]ָהּ (her)
[PP]
[pp]
[pp]
[prep] לְ (to)
[np]
[pron]ְךָ (you)

3.3.4 Object clauses

Example 14 (from Genesis 3:11): The object clause as a whole is the object of the main verb:

[CL]
[S]
[np]
[pron] מִי (who?)
[V]
[vp]
[verb] הִגִּיד (he told)
[PP]
[pp]

[pp]
[prep] לְ (to)
[np]
[pron]ךָ֫ (you)
[O]
[CL]
[cjp]
[cj] כִּי (that)
[CL]
[P]
[adjp]
[adj] עֵירֹם (naked)
[S]
[np]
[pron] אַתָּה (you))

3.4 Phrase-level Attributive Modifiers

There are various phrase-level attributive modifiers. These attributively specify or delimit the meaning of their head. They include adjectives, numerals adverbs, definite articles, participles, demonstrative pronouns, prepositional phrases, nouns in apposition, nouns in construct relation, and relative clauses. Initially, phrase level attributive modifiers were simply allowed to combine with the nominals they appeared to modify. So, preliminary rule names are often just descriptive of what two elements were combined together.

3.4.1 Adjectives or Numerals

Adjectives that modify a noun are shown as forming a larger noun phrase with the noun The rule name is NpAdjp. (Numerals act like adjectives—the corresponding rule name is NpNump, see for example Genesis 1:31. Both adjectives and numerals can function as nouns too.) Example 15 (from Genesis 21:8):

[cjp]
[cj] וְ (and)
[CL]
[V]
[vp]
[verb] יַּעַשׂ (he held)
[S]
[np]
[noun] אַבְרָהָם (Abraham)
[O]
[np]
[np]
[noun] מִשְׁתֶּה (feast)
[adjp]
[adj] גָּדוֹל (great)

3.4.2 Adverbs

Example 16 (from Genesis 1:31): The adverb modifies an adjective to form a larger adjectival phrase:

[CL]
[P]
[adjp]
[adjp]
[adj] טוֹב (good)

[advp]
[adv] מְאֹד (very)

3.4.3 Definite articles

Example 17 (also from Genesis 21:8): The definite article modifies a noun to form a larger noun phrase with the noun (rule name DetNP):

[cjp]
[cj] וְ (and)
[CL]
[V]
[vp]
[verb] יִגְדַּל (he grew)
[S]
[np]
[art] הַ (the)
[np]
[noun] יֶלֶד (child)

3.4.4 Participles

Participles (technically participial clauses as even single participles are automatically promoted to a clause) that modify a noun are shown as forming a larger noun phrase with the noun. After the participial clause is promoted to an adjp, it combines with the np in the same ways as adjectives. Example 18 (from Genesis 24:43):

[CL]
[np]
[art] הָ (the)
[np]
[np]
[noun] עַלְמָה (maiden)
[adjp]
[art]
[art] הַ (the)
[adjp]
[CL]
[V]
[vp]
[verb] יֹצֵאת (one coming)
[PP]
[pp]
[pp]
[prep] לְ (to)
[CL]
[V]
[vp]
[verb] שְׁאֹב (draw)

3.4.5 Demonstrative pronouns

Example 19 (from Genesis 7:1): Demonstrative pronouns function just like adjectives:

[np]
[art] (the)
[np]

[np]
[noun] דּוֹר (generation)
[adjp]
[art] הַ (the)
[adjp]
[adj] :זֶּה (this)

3.4.6 Prepositional phrases

Prepositional phrases that modify a noun form a larger noun phrase by the descriptively-named Np-Pp rule. Example 20 (from Genesis 3:6):

[cjp]
[cj] וְ (and)
[CL]
[V]
[vp]
[verb] תִּתֶּן (she gave)
[PP]
[pp]
[cjp]
[cj] גַּם־ (also)
[pp]
[pp]
[prep] לְ (to)
[np]
[np]
[np]
[noun] אִישׁ (husband of)
[np]
[pron]ָהּ (her)
[pp]
[pp]
[prep] עִמּ (with)
[np]
[pron]ָהּ (her)

3.4.7 Apposition

In these cases, the nominal in apposition functions similarly to an adjective, but the form is that of two nominals in apposition. So, the two nominals are represented as nominals forming a larger np by the Np-Appos rule. Example 21 (from Genesis 20:2):

[cjp]
[cj] וְ (and)
[CL]
[V]
[vp]
[verb] יֹּאמֶר (said)
[S]
[np]
[noun] אַבְרָהָם (Abraham)
[PP]
[pp]
[pp]

[prep] אֶל־ (about)
[np]
[np]
[noun] שָׂרָה (Sarah)
[np]
[np]
[noun] אִשְׁתּ (wife of)
[np]
[pron]וֹ (him)

3.4.8 Construct relation

The construct state grammatically links a word to the following word(s), forming a single grammatical unit called a construct chain (Rule name NPofNP). The qualified precedes the qualifier. Literal English translations often translate words in a construct chain with the word “of” between them. Example 22 (from Genesis 41:10):

[np]
[np]
[noun] מִשְׁמַר (custody of)
[np]
[np]
[noun] בֵּית (house of)
[np]
[np]
[noun] שַׂר (captain of)
[np]
[art] הַ (the)
[np]
[noun] טַבָּחִים (guards)

3.4.9 Relative clauses

Relative clauses are represented as clauses modifying and forming larger noun phrases together with the noun phrases they modify by the descriptively-named rule Np-Relp. Example 23 (from Genesis 3:3):

[np]
[np]
[noun] פְּרִי (fruit of)
[np]
[art] הָ (the)
[np]
[np]
[noun] עֵץ (tree)
[relp]
[rel] אֲשֶׁר (that)
[CL]
[P]
[pp]
[pp]
[prep] בְּ (in)
[np]
[np]
[noun] תּוֹךְ־ (middle of)
[np]

[art] הַ (the)
[np]
[noun] גָּן (garden)

3.5 Adverbial Function (ADV)
Adverbials Function (ADV) further specifies the circumstances under which a verb takes place. The trees analysis was previously done with a broader definition of adverbs than OSHB uses, including adverbs proper, modal words, and some particles (including negatives). These will require reevaluation.

3.5.1 Adverbs
Example 24 (Genesis 15:5): The demonstrative adverb "so" refers back to given information that describes the nature of the process or events:

[CL]
[ADV]
[advp]
[adv] כָּה (so)
[V]
[vp]
[verb] יִהְיֶה (he shall be)
[S]
[np]
[np]
[noun] זַרְע (offspring of)
[np]
[pron] :ךָ(you)

3.5.2 Modal words
Example 25 (from Genesis 17:19): The modal word indicates the denial of an expectation or view that the speaker thinks a listener holds:

[CL]
[ADV]
[advp]
[adv] אֲבָל (yes but)
[S]
[np]
[np]
[noun] שָׂרָה (Sarah)
[np]
[np]
[noun] אֵשֶׁת (wife of)
[np]
[pron]ךָ֣ (you)
[V]
[vp]
[verb] יֹלֶדֶת (bearing)
[PP]
[pp]
[pp]
[prep] ל (for)
[np]
[pron]ךָ֣ (you)

[O]
[np]
[noun] בֵּן (a son)

3.5.3 Negatives
Example 26 (from Genesis 2:17): The negative particle negates a command:
[CL]
[ADV]
[advp]
[adv] לֹא (not)
[V]
[vp]
[verb] תֹּאכַל (you must eat)
[PP]
[pp]
[pp]
[prep] מִמּ (from)
[np]
[pron]ֶּנּוּ (him)

3.5.4 Infinitive absolutes
Example 27 (from Genesis 2:16): The nature or scope of the verbal idea is defined more clearly:
[CL]
[PP]
[pp]
[pp]
[prep] מִ (from)
[np]
[np]
[noun] כָּל (any of)
[np]
[np]
[noun] עֵץ־ (tree of)
[np]
[art] הַ (the)
[np]
[noun] גָּן (garden)
[ADV]
[advp]
[verb] אָכֹל (to eat [as you please)
[V]
[vp]
[verb] :תֹּאכֵל (you may eat)

3.5.5 Verb + infinitive absolute + infinitive absolute
Example 28 (from Genesis 8:7): The main verb involves motion to a certain place. The first infinitive absolute repeats the stem of the main verb while the second infinitive absolute involves a second action that occurs simultaneously with the first:
[cjp]
[cj] וְ (and)
[CL]

[V]
[vp]
[verb] יֵּצֵא (it [the raven] went)
[ADV]
[advp]
[advp]
[verb] יָצוֹא (to go out)
[cjp]
[cj] וְ (and)
[advp]
[verb] שׁוֹב (to return)

3.6 Prepositional Phrase Function (PP)
In the current trees prepositional phrases that do not modify a nominal are consistently labeled with prepositional phrase function (PP) at the clause level. Our partner the Groves Center had specifically requested this early in our partnership to distinguish prepositional phrases from adverbs. More evaluation is needed to determine whether prepositional phrases should be grouped with adverbials as ADV to conform to the Greek treebank. Among the considerations that need to be made are whether to annotate prepositional phrases that seem to function as complements differently from the ones that function strictly as adverbial modifiers.

3.6.1 Prepositional phrases as adverbial modifiers
Example 29 (from Genesis 31:54): The prepositional phrase functions as an adverbial modifier of where the action takes place:

[cjp]
[cj] וַ (and)
[CL]
[V]
[vp]
[verb] יִּזְבַּח (he offered)
[S]
[np]
[noun] יַעֲקֹב (Jacob)
[O]
[np]
[noun] זֶבַח (sacrifice)
[PP]
[pp]
[pp]
[prep] בְּ (in)
[np]
[art] (the)
[np]
[noun] הָר (hill country)

3.6.2 Prepositional phrases as complements
Example 30 (from Genesis 12:7): The prepositional phrase seems to mark the indirect object:

[CL]
[PP]
[pp]
[pp]

[prep] לְ (to)
[np]
[np]
[noun] זֶרַע (offspring of)
[np]
[pron]ךָּ֖ (you)
[V]
[vp]
[verb] אֶתֵּן (I will give)
[O]
[np]
[omp]
[om] אֶת־ (-)
[np]
[art] הָ (the)
[np]
[np]
[noun] אֶרֶץ (land)
[adjp]
[art] הַ (the)
[adjp]
[adj] זֹּאת (this)

3.7 Predicate (P)
Predicate (P) usually marks predicate nominals. Nouns, adjectives, some pronouns, and prepositional phrases commonly serve as predicates in nominal clauses. In the current trees, unless a clause without a verb is seen as having an elided verb (V), it will have P as the core constituent.

3.7.1 Nouns as predicates
Example 31 (from Genesis 45:3): The predicate of the nominal clause is a noun:
[CL]
[S]
[np]
[pron] אֲנִי (I)
[P]
[np]
[noun] יוֹסֵף (Joseph)

3.7.2 Adjectives as predicates
Example 32 (from Genesis 6:2): The predicate of the nominal clause is an adjective:
[CL]
[P]
[adjp]
[adj] טֹבֹת (beautiful ones)
[S]
[np]
[pron] הֵנָּה (they)

3.7.3 Interrogative pronouns as predicate
Example 33 (from Genesis 48:8): The predicate of the nominal clause is an interrogative pronoun (note that OSHB parses this as an interrogative particle):

[CL]
[P]
[np]
[pron] מִי־ (who?)
[S]
[np]
[adjp]
[adj] אֵלֶּה: (these)

3.7.4 Prepositional phrases as predicate
Example 34 (from Genesis 3:16): The predicate of the nominal clause is the prepositional phrase:

[CL]
[P]
[pp]
[pp]
[prep] אֶל־ (to)
[np]
[np]
[noun] אִישׁ (husband of)
[np]
[pron]ךָ֔ (you)
[S]
[np]
[np]
[noun] תְּשֻׁוקָ֖ת (desire of)
[np]
[pron]ךְ֑ (you)

3.8 Second Object (O2)
Some verbs take two objects. There are two main types. The first type involves an object of person (the first object) and an object of thing (the second object). The second type also a direct object and an object complement. In the current version of the trees, these two types are not yet distinguished. The distinguishing of object complements (OC) is planned for the near future.

3.8.1 Object of Person and Object of Thing
Example 35 (from Genesis 37:23):

[CL]
[V]
[vp]
[verb] יַפְשִׁיטוּ (they stripped)
[O]
[np]
[omp]
[om] אֶת־ (-)
[np]
[noun] יוֹסֵף (Joseph)
[O2]
[np]
[np]
[omp]
[om] אֶת־ (-)

[np]
[np]
[noun] כֻּתֶּנֶת (robe of)
[np]
[pron]וֹ (him)

3.8.2 Direct Object and Object Complement
Example 36 (from Genesis 30:6):

[CL]
[V]
[vp]
[verb] קָרְאָה (she called)
[O]
[np]
[np]
[noun] שְׁמ (name of)
[np]
[pron]וֹ (him)
[O2]
[np]
[noun] :דָּן (Dan)

Example 37 (from Genesis 37:23):

[CL]
[V]
[vp]
[verb] אֶעֱשֶׂה (I can prepare)
[O]
[np]
[omp]
[om] אֹת ()
[np]
[pron]ָם (them [choice young goats])
[O2]
[np]
[noun] מַטְעַמִּים (tasty foods)

3.9 “Bridge” Structures
Conjunctions (both coordinating and subordinating) and prepositions function as “bridges” between their children and their own heads. Words in apposition are covered under 3.4.6 Phrase-level attributive modifiers.

3.9.1 Coordinating conjunctions (conj)
Coordinating conjunctions simply connect other structures together (the first item is usually arbitrarily assigned as the head of the resultant overall coordinated structure). Coordinators may connect two or more single words, phrases, or clauses. The current version of the trees does not attempt to connect clauses. Example 38 (from Genesis 12:16): Seven noun phrases are coordinated by six coordinating conjunctions:

[np]
[np]
[noun] צֹאן־ (sheep)

[cjp]
[cj] וְ (and)
[np]
[noun] בָּקָר (cattle)
[cjp]
[cj] וְ (and)
[np]
[noun] חֲמֹרִים (donkeys)
[cjp]
[cj] וְ (and)
[np]
[noun] עֲבָדִים (male servants)
[cjp]
[cj] וְ (and)
[np]
[noun] שְׁפָחֹת (female servants)
[cjp]
[cj] וְ (and)
[np]
[noun] אֲתֹנֹת (female donkeys)
[cjp]
[cj] וְ (and)
[np]
[noun] גְּמַלִּים׃ (camels)

Example 39 (from Genesis 12:1): Three prepositional phrases are coordinated by two coordinating conjunctions:

[CL]
[V]
[vp]
[verb] לֶךְ־ (leave!)
[PP]
[pp]
[pp]
[prep] לְ (for)
[np]
[pron]ךָ֣ (you)
[PP]
[pp]
[pp]
[pp]
[prep] מִ (from)
[np]
[np]
[noun] אַרְצ (country of)
[np]
[pron]ךָ֣ (you)
[cjp]
[cj] וְ (and)
[pp]
[pp]

[prep] מִ (from)
[np]
[np]
[noun] מְוֹלַדְתְּ (people of)
[np]
[pron]ךָ֫ (you)
[cjp]
[cj] וְ (and)
[pp]
[pp]
[prep] מִ (from)
[np]
[np]
[noun] בֵּית (house of)
[np]
[np]
[noun] אֲבִי (father of)
[np]
[pron] הּ (you)

3.9.2 Prepositions (prep)

Technically the preposition governs the nominal that is its object. However, prepositions are regarded as function words and the more salient relationship is between object of the preposition and the verb. So, the object of the preposition is the head of prepositional phrases formed by the PrepNp rule (even though the preposition could have been marked as the head instead). Example 40 (from Genesis 9:11):

[CL]
[ADV]
[advp]
[adv] לֹא־ (not)
[V]
[vp]
[verb] יִכָּרֵת (will be cut off)
[S]
[np]
[np]
[noun] כָּל־ (all of)
[np]
[noun] בָּשָׂר (life)
[ADV]
[advp]
[adv] עוֹד (ever)
[PP]
[pp]
[pp]
[prep] מִ (by)
[np]
[np]
[noun] מֵי (waters of)
[np]
[art] הַ (the)
[np]

[noun] מַבּוּל (flood)

We often find an infinitive construct governed by a preposition. Frequently, it is equivalent to a finite tense preceded by a conjunction. Example 41 (from Genesis 27:45):

[PP]
[pp]
[pp]
[prep] עַד־ (when)
[CL]
[V]
[vp]
[verb] שׁוּב (subsides)
[S]
[np]
[np]
[noun] אַף־ (anger of)
[np]
[np]
[noun] אָחִי (brother of)
[np]
[pron] דְּ (you)
[PP]
[pp]
[pp]
[prep] מִמ (from)
[np]
[pron]ךָּ (you)

Example 42 (from Genesis 31:19): With the preposition ל, it often indicates the direction, aim, or purpose of an action:

[CL]
[S]
[np]
[noun] לָבָן (Laban)
[V]
[vp]
[verb] הָלַךְ (went)
[PP]
[pp]
[pp]
[prep] לְ (to)
[CL]
[V]
[vp]
[verb] גָּזַז (shear)
[O]
[np]
[omp]
[om] אֶת־ (-)
[np]
[np]

[noun] צֹאן (sheep of)
[np]
[pron]וֹ (him)

3.9.3. Subordinating conjunctions

Like with prepositions, while subordinating conjunctions could be seen as the heads of subordinate clauses, they are viewed as “bridges” since they are function words and the actual relationship is between the subordinate clause and the main clause. Example 43 (from Genesis 18:21): The conditional subordinate clause precedes the main clause:

[CL]
[CL]
[cjp]
[cj] אִם־ (if)
[CL]
[ADV]
[advp]
[adv] לֹא (not)
[CL]
[V]
[vp]
[verb] :אֵדָעָה (I will know)

Example 44 (from Genesis 27:4): Many prepositions serve as subordinating conjunctions with two in succession in this example:

[CL]
[CL]
[V]
[vp]
[verb] אֹכֵלָה (I may eat)
[CL]
[pp]
[pp]
[prep] בַּ ()
[np]
[noun] עֲבוּר (so that)
[CL]
[CL]
[V]
[vp]
[verb] תְּבָרֶכ (she may bless)
[O]
[np]
[pron]ךָ (you)
[S]
[np]
[np]
[noun] נַפְשׁ (self of)
[np]
[pron]ִי (me)
[CL]
[pp]

[pp]
[prep] בְּ ()
[advp]
[adv] טֶרֶם (before)
[CL]
[V]
[vp]
[verb] :אָמוּת (I die)

4 How Specific Constructions Are Annotated

4.1 Ellipsis

While the phenomenon of verb ellipsis is recognized, its presence is currently marked only by the assignment of Verb Elided Clause as the ClType. (Shared objects are also often elided from one clause to the next and they are likewise not marked in the current trees.) Other constituents of the clause, such as ADV, would receive their normal labels as if the V were present (even though it is not). Example 45 (from Genesis 18:21): The verb "have done" is elided and assumed to be carried over from the previous clause:

[CL]
[CL]
[cjp]
[cj] אִם־ (if)
[CL]
[ADV]
[advp]
[adv] לֹא (not)
[CL]
[V]
[vp]
[verb] :אֵדָעָה (I will know)

4.2 Particles

Particles belong under a morphological category of uninflected function words. A number of OSHB particle types currently are represented generally as particles while others (e.g., object marker, definite article, relative particle) have been specifically named. Particles are typically invariable in form and some partially overlap with the broader definition of adverbs that modify at the clause or sentence level.

4.3 Direct Speech

Clauses of direct speech are left for now as their own separate clauses and are not connected with the clause that introduced the speech.