

Tree-Based Approaches to Biblical Texts

Andi Wu & Randall Tan
Asia Bible Society / GrapeCity

Trees and Treebanks

- Tree:

A representation of the syntactic/grammatical structure of a sentence in the form of a tree diagram

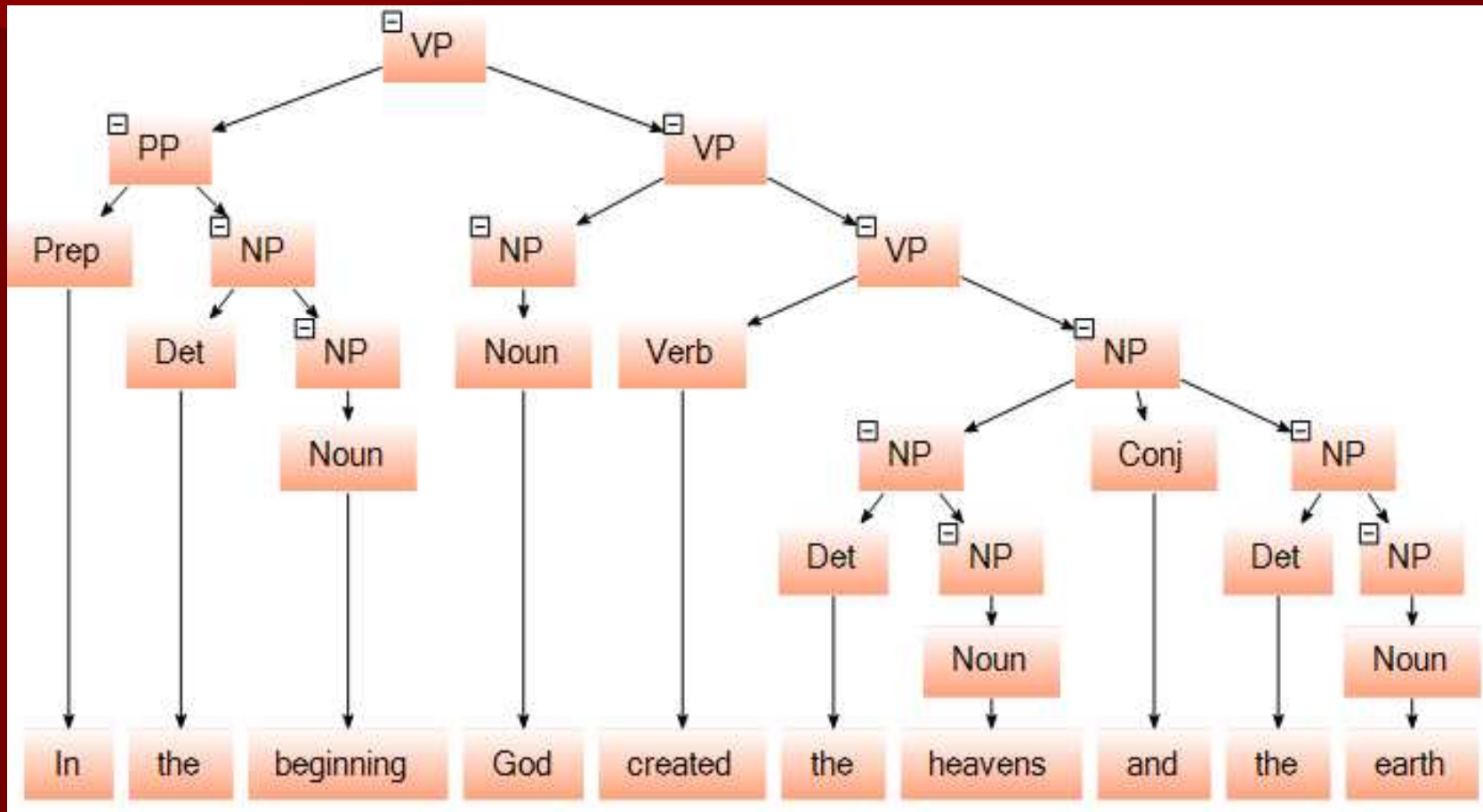
- Treebank:

The collection of trees for a particular text:

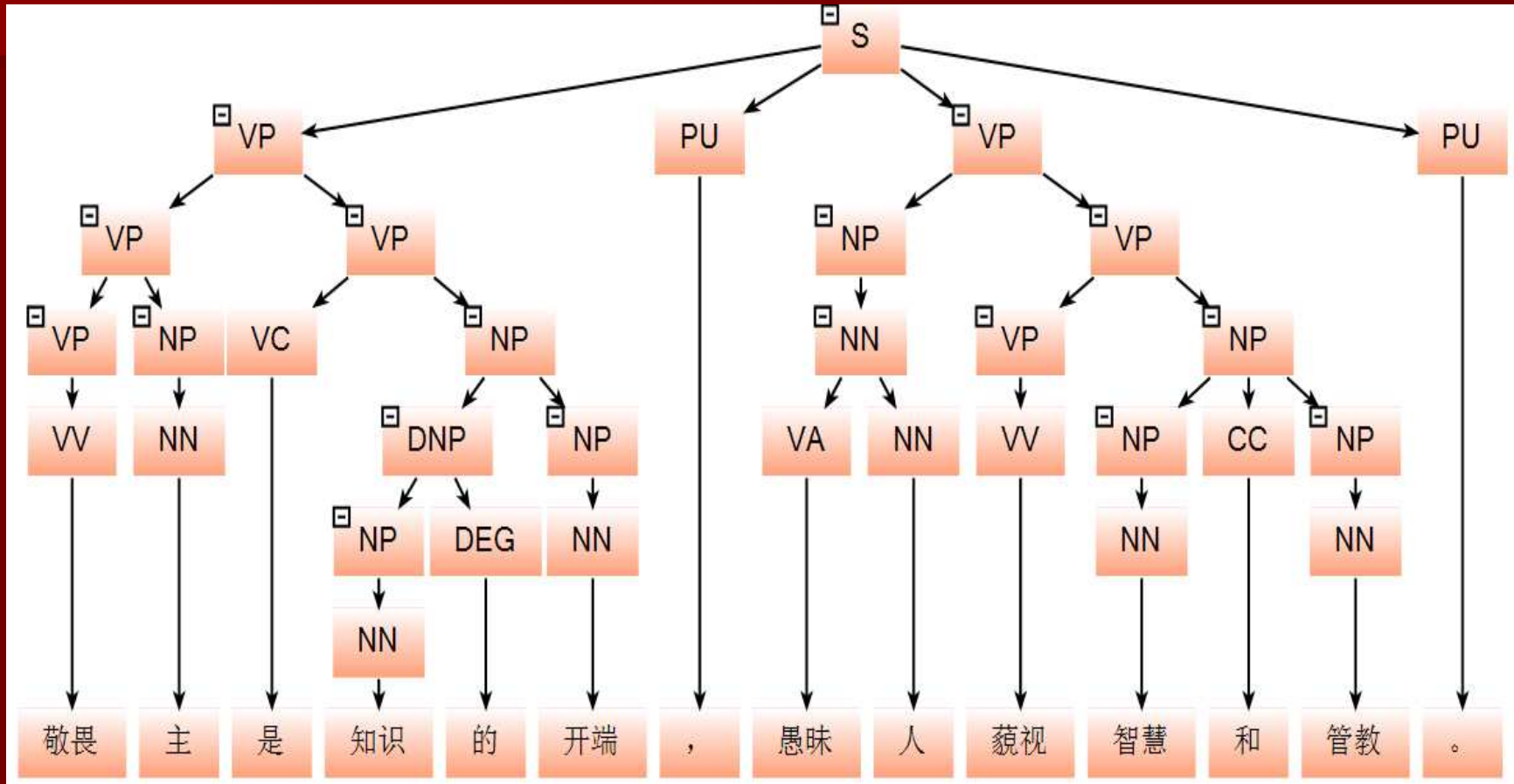
- Hebrew treebank of the Old Testament
- Greek treebank of the New Testament
- Chinese treebank of a Chinese translation of the Bible
- English treebank of an English translation of the Bible



An English Tree



A Chinese Tree



Tree-Based Applications

Supporting the CSB translation project:

- Tree alignment
- Tree-based translation memory
- Tree-based concordance
- Tree-based interlinear view
- Probabilistic Hebrew Synonym Finder
- Probabilistic Similar Verse Finder

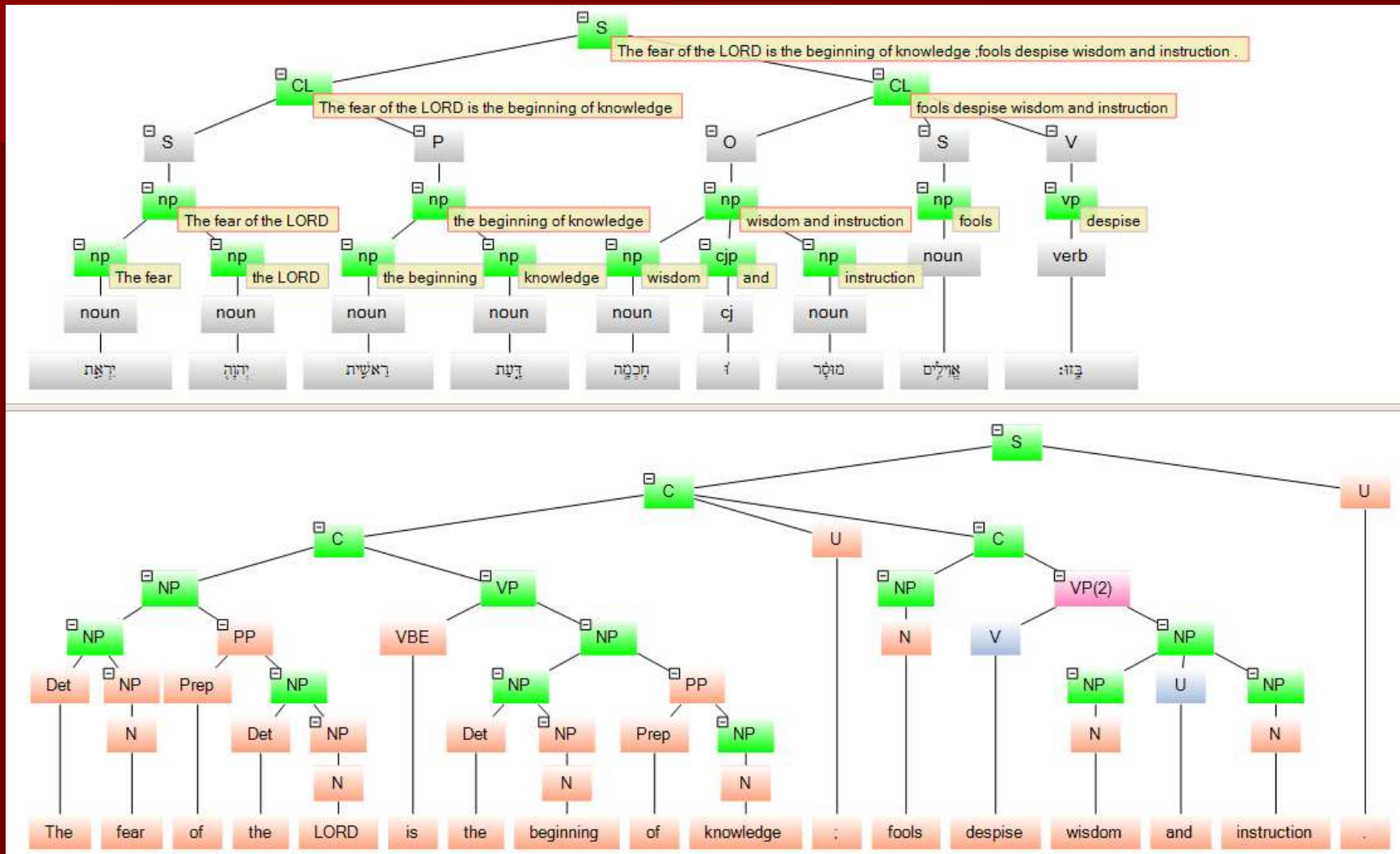


Tree Alignment

- Node linking between source languages and target languages
- The linked node can be a word, a phrase, or a clause
- Word-to-phrase, phrase-to-word
- Trees make automatic alignment more accurate
- A tool for checking the accuracy of Bible translation
- An objective way of measuring how literal/dynamic the translation is
- Linked data enables many other applications



Aligned Hebrew and English Trees



Tree-based Translation Memory

- Traditional translation memory:
 - Word-based
- Tree-based translation memory:
 - Chunk-based
(a chunk can be word, a phrase, or clause)
 - Relation-based
(pairs of words in certain grammatical relations)



Tree-based Translation Memory

■ Chunks

The fear of the LORD is the beginning of knowledge ;
fools despise wisdom and discipline.

...

the fear of the LORD
the beginning of knowledge
wisdom and discipline
despise wisdom and discipline

...



Tree-based Translation Memory

■ Head-Modifier Pairs

The fear of the LORD is the beginning of knowledge ;
fools despise wisdom and discipline.

fear-LORD

beginning-knowledge

fool-despise

despise-wisdom

despise-discipline

wisdom-discipline

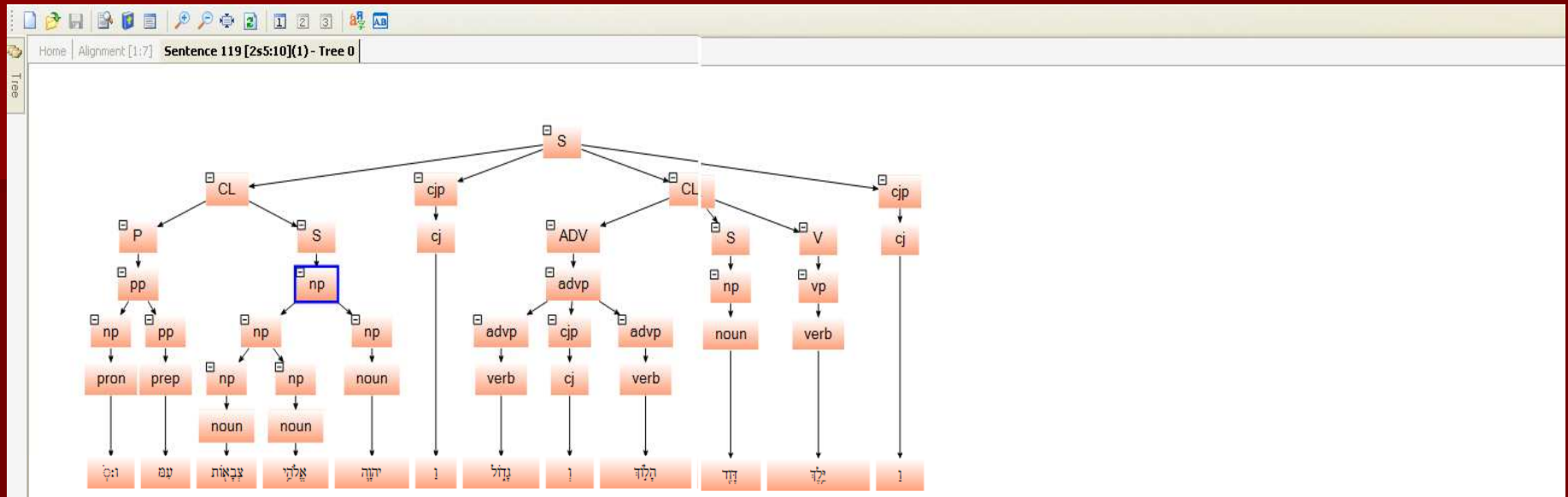


Tree-based Translation Memory

- The translators/editors have access to all the existing translations for a given word, phrase or clause.
- Tree-Based Translation: getting an overview of the translations for a chunk or relational pair
- Tree-Based Concordance: viewing the chunks and pairs in their contexts.
- Demo



Concordance of a chunk



Concordance

Search: 3068 0430 6635

Go [Export to Excel](#)

Strong Numbers Lemma

			呢?
am4:13	那创世造风, 将心意指示人, 使晨光变为幽暗, 脚踏在地之高处的, 他的名是	耶和华万军之神	!
am5:14	你们要求善, 不要求恶, 就必存活。这样,	耶和华万军之神	必照你们所说的, 与你们同在。
am5:15	要恶恶好善, 在城门口秉公行义, 或者	耶和华万军之神	向约瑟的余民施恩。
am5:16	主	耶和华万军之神	如此说: “在一切宽阔处必有哀号的声音, 在各街市上必有人说: ‘哀哉! 哀哉!’ 又必叫农夫来哭哀歌的来举哀。”
am6:8	主耶和华	万军之神	指着自已起誓说: “我憎恶雅各的荣华, 厌弃他的宫殿, 因此我必将城和其中所有的都交付敌人。”
HCSBX			
2s5:10	David became more and more powerful, and the	LORD God of Hosts	was with him .
1k19:10	¶ He replied, “I have been very zealous for the	LORD God of Hosts	, but the Israelites have abandoned Your covenant , torn down Your altars , and killed Your prophets with the sword left , and they are looking for me to take my life .”
1k19:14	¶ “I have been very zealous for the	LORD God of Hosts	,” he replied, “ but the Israelites have abandoned Your covenant , torn down Your altars , and killed Your prophets with sword . I alone am left , and they’re looking for me to take my life .”
ps59:6	¶	LORD God of Hosts	, You are the God of Israel , ¶ rise up to punish all the nations ; ¶ do not show grace to any wicked traitors . • Selah
ps80:5	¶	LORD God of Hosts	, ¶ how long will You be angry ¶ with Your people’s prayers ?



Translations of a collocation

Home | Alignment [1:7] | Sentence 4 [gn26:5](1) - Tree 0

Translations

Search 8451+8104 Exact Match Strong Numbers Lemma

csv
 遵守+律法(7) 遵守+教训(1) 谨守+律法(1)

cuv
 守+律法(4) 遵守+律法(4) 遵行+律法(1) 遵守+法度(1)

eng
 keeping+law(2) I will obey+law(1) keep+instructions(1) obey+law(1) I will keep+law(1)
 keep+law(1) kept+law(1) kept+laws(1) keep+laws(1)

hcsbx
 keep+law(2) obey+instruction(2) observe+laws(1) listens to+instruction(1)
 follow+instruction(1) kept+instructions(1) keep+instruction(1) keep+instructions(1)

nkjv
 I keep+law(2) kept+laws(1) keep+law(1) keep+laws(1) who keeps+law(1) [men] do
 keep+law(1) kept+law(1) such as keep+law(1) They shall keep+laws(1)

pth
 谨守+律法(2) 遵行+律法(1) 守+律法(1) 遵从+教导(1)

sgs
 遵守+法律(4) 遵循+法律(1) 遵守+法典(1) 守+法律(1) 遵守+法令(1) 遵守+法规(1)



Tree-based Interlinear

- Dynamic linking
- Zoom in or out to view correspondences at different granularity
- Demo



Tree-based Interlinear

Reference extension	
Pro 1:2	
BHS	לְ דַעַת חִכְמָה וּמוֹסֵר לְ הַבֵּן אִמְרֵי בִינָה:
CSB	为要明晓智慧和教训，领悟含义深刻的格言；
CUV	要使人晓得智慧和训诲，分辨通达的言语，
HCSB	For learning what wisdom and discipline are ; for understanding insightful sayings ;

Reference extension	
Pro 1:2	
BHS	לְ דַעַת חִכְמָה וּמוֹסֵר לְ הַבֵּן אִמְרֵי בִינָה:
CSB	为要明晓智慧和教训，领悟含义深刻的格言；
CUV	要使人晓得智慧和训诲，分辨通达的言语，
HCSB	For learning what wisdom and discipline are ; for understanding insightful sayings ;

Reference extension	
Pro 1:2	
BHS	לְ דַעַת חִכְמָה וּמוֹסֵר לְ הַבֵּן אִמְרֵי בִינָה:
CSB	为要明晓智慧和教训，领悟含义深刻的格言；
CUV	要使人晓得智慧和训诲，分辨通达的言语，
HCSB	For learning what wisdom and discipline are ; for understanding insightful sayings ;



Probabilistic Hebrew Synonym Finder

Why build a probabilistic synonym finder automatically?

- Translators/editors need to know the synonyms to better align target words to source words
- Existing synonym dictionaries are incomplete
- Manual construction of the knowledge base is time-consuming
- Synonym sets are not discrete
- Large quantities of linked data to be leveraged.



Hebrew Synonym Finder

- Defining synonyms in terms of semantic space:

Two words are synonymous when their semantic spaces overlap.

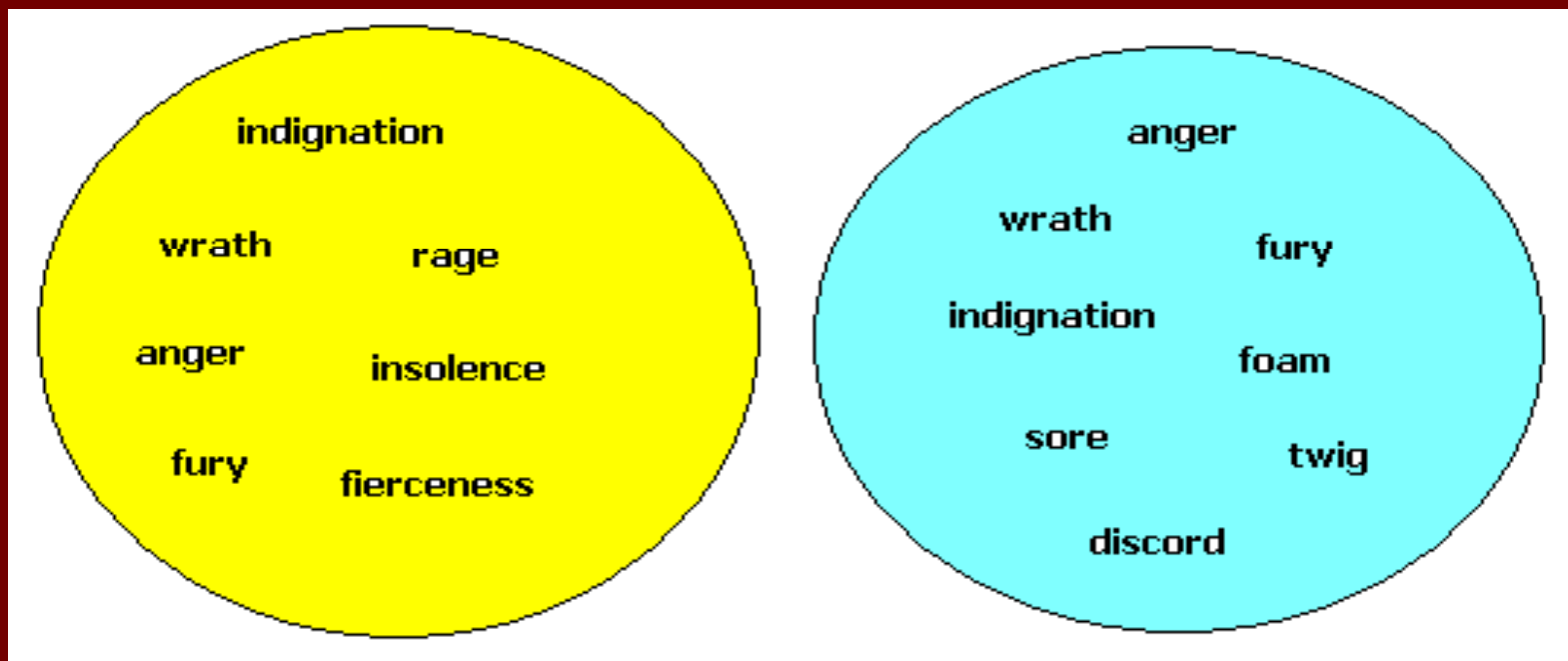
- Defining the semantic space of Hebrew words:

The semantic space of a Hebrew word consists of all the words that have been used to translate this word in other languages.



Hebrew Synonym Finder

Semantic spaces of זַעַם and קֶצֶף



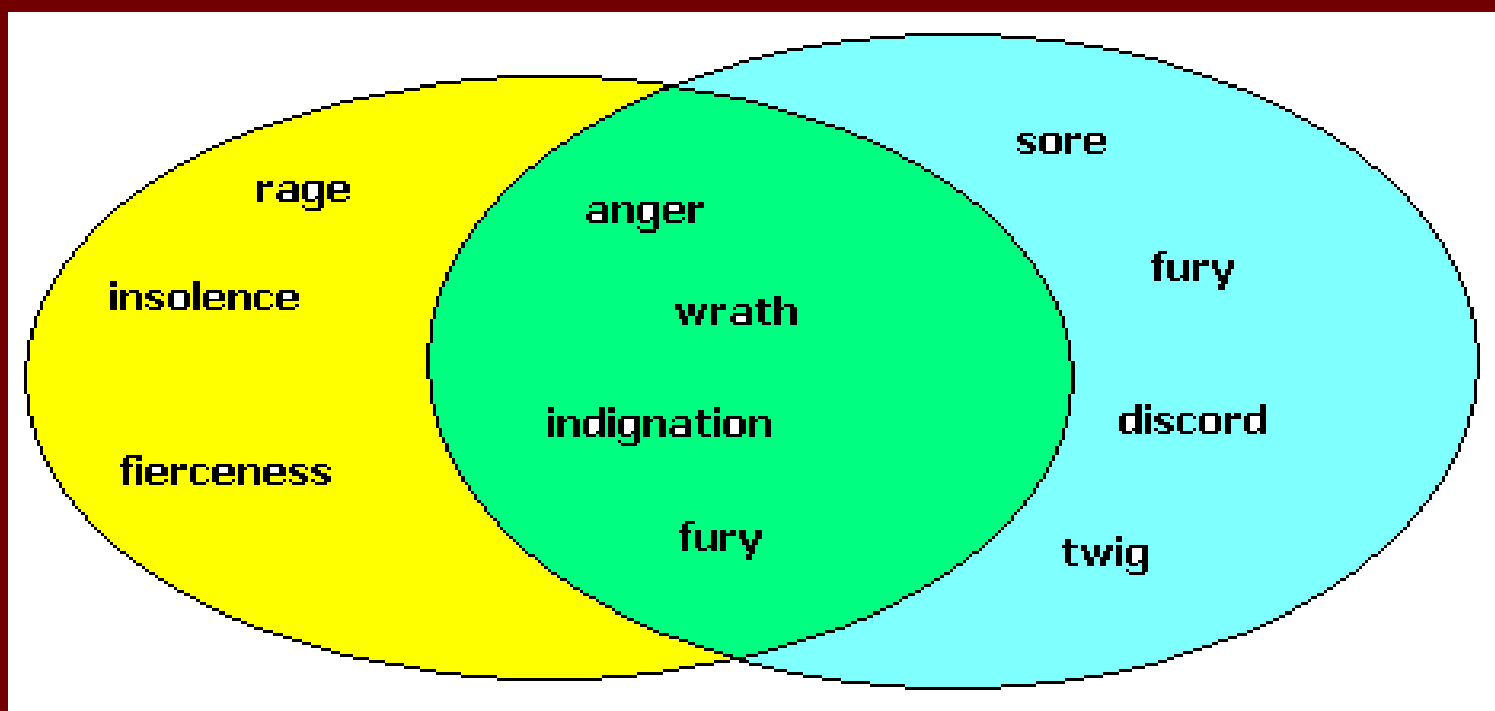
זַעַם

קֶצֶף



Hebrew Synonym Finder

Overlap between two semantic spaces



זַעַם

קִצְף



Hebrew Synonym Finder

Degree of synonymy:

intersection of sets / union of two sets

Score:

joint probabilities of several different ways of computing the intersection

Parameters to tune:

- Chinese vs. English
- Set of words vs. bag of words
- Words vs. characters

Demo



Hebrew Synonym Finder

Search Term
 Hebrew Word: 0998 Strong Lemma Unicode Lemma

Range: From 18 to 22 **Details**

POS: Noun Count: 23 Strong: 0998 Lemma: B.IYNFH Unicode: גִּיּוֹן

English: understanding(1284) insight(5) wisdom(5) own understanding(4) [is] understanding(4) have unc

Chinese: 聪明(549) 悟性(283) 明达(241) 明智(52) 哲理(50) 见识(18) 智慧(10) 明哲(6) 知识(5) 睿智(5)

Scope: _____ **Get Synonyms**

Synonyms	Distance	English	Chinese
7922 (8) Noun שָׂכַל	104.17192061	insight(40) understanding(29) wisdom(18)	见识(37) 明慧(25) 聪明(20)
8394 (27) Noun תְּבוּנָה	107.73836566	sense(1) esteem(1)	智慧(17) 明智(5)
0995 (83) Verb בִּין	124.77622320	The discretion(1)	明白(2)
1847 (61) Noun דַּעַת	127.34964455	Understanding [is](1)	明理(2)
7919 (27) Verb שָׂכַל	136.17474194	high regard(1)	识见(1)
2940 (6) Noun טַעַם	141.38463669	name(1)	明哲的(1)
2451 (91) Noun חֲכָמָה	149.15784361	A understanding(1)	高见(1)
3820 (263) Noun לֵב	156.42698138		智识(1)
1844 (2) Noun דָּשָׂה	156.81062611		名(1)
8454 (10) Noun תּוֹשָׁיָה	158.71318601		明智的(1)
2454 (4) Noun חֲכָמָה וְחָכְמָה	160.47211648		通达事理(1)
3824 (47) Noun לֵבָב	160.90354046		良好的教养(1)
1843 (5) Noun דָּעַת	168.98142472		睿智颖悟(1)
4093 (1) Noun טָעַת	172.42048310		
2449 (23) Verb חָכַם	179.81200914		

Showing the top 50% of 31 records



Probabilistic Similar Verse Finder

- Similar verses:
 - Verses containing similar clauses
- Similar clauses:
 - Clauses which are not necessarily identical but have more or less the same meaning
- A supplement to the translation memory
 - Translation memory: exact search
 - Similar verse finder: fuzzy search



Similar Verse Finder

- Defining similar clauses in terms of their semantic spaces:

Two clauses are similar when their semantic spaces have a significant overlap.

- Defining the semantic space of a clause:

The semantic space of a clause consists of all the head-modifier pairs found in the tree of this clause.



Similar Verse Finder

A: Ps19:1 Better a poor man who lives with integrity than someone who has deceitful lips and is a fool .

B: Ps28:6 Better a poor man who lives with integrity than a rich man who distorts right and wrong .

A & B

better-poor

better-than

poor-walking

walking-integrity

integrity-him

than-perverse

A only

perverse-lips

lips-him

he-fool

perverse-fool

B only

perverse-ways

he-rich

perverse-rich



Similar Verse Finder

Degree of similarity:

intersection of sets / union of two sets

Score:

joint probabilities of several different ways of computing the intersection

Parameters to tune:

- Head-modifier pairs to be included in the set
- Labeled vs. unlabeled pairs
- Clause-based vs. verse-based

Demo



Find Similar Verses

Book: Psalms
 Chapter: 14
 Verse: 1

Show Verse Text
Get Similar Verses

Manuscript
 למנצח לְדָוִד לְמַרְגָּלָה לְבוֹאֵי יָסָה שְׁחִיתוּהָת עִיבוּעַל־יְהוָה אֱיִן אֱמֶשָׁה-טוֹב:

Translations

CSB 大卫的诗，交给音乐总监。愚昧人心里说：“没有神！”他们败坏，行事可憎，无人行善。

CUV 大卫的诗，交于伶长。愚顽人心里说：“没有神！”他们都是邪恶，行了可憎恶的事，没有一个人行善。

HCSB ¶ For the choir director . Davidic . ¶ The fool says in his heart , “ God does not exist . ” ¶ They are corrupt ; they do vile deeds . ¶ There

Similar Verses

Verses	Difference
ps53:2	1.72969
ps10:13	2.93857
1k12:26	3.12089
ps10:11	3.12089
ps10:6	3.20093
ps14:3	3.74950
pr6:14	3.74950
ps53:4	3.74950
gn17:17	3.89408
gn27:41	4.00328

Manuscript
 נְבִיאַ מַרְגָּלָה לְבוֹאֵי יָסָה שְׁחִיתוּהָת עִיבוּעַל־יְהוָה אֱיִן אֱמֶשָׁה-טוֹב:

CSB 大卫的训诲诗，交给音乐总监，用“麻哈拉”的曲调。愚昧人心里说：“没有神！”他们败坏，行事可憎，无人行善。

CUV 大卫的训诲诗，交于伶长。调用麻哈拉。愚顽人心里说：“没有神！”他们都是邪恶，行了可憎恶的罪孽，没有一个人行

HCSB ¶ For the choir director : on Mahalath . A Davidic • Maskil . ¶ The fool says in his heart , “ God does not exist . ” ¶ They are corrupt .



Summary

Syntactically annotated Biblical texts (trees) provide the foundation for many intelligent applications in the study and translation of the Bible.

