

New Technology for Bible Understanding and Translation

Sean Boisen

Clear Bible, Inc. (<https://www.clear.bible/>)



<https://bit.ly/sboisen-newtech>



About Me

- MA in Computational Linguistics
- Technologist, information innovator
- 19 years with BBN Technologies in Human Language Technology
 - Researcher, technology manager, technology transition
 - Gov't funded R&D in named entity extraction, information extraction, speech and language processing
- 15+ years with Faithlife/Logos Bible Software
 - Senior Information Architect, Director of Content Innovation
 - Building and managing biblical knowledge
- 3 months (!) with Clear Bible
 - Principal Data Scientist
 - Product Lead for Enhanced Resource++

Outline

- Logos: Biblical Places dataset
- Logos: Atlas, Biblical Places
- Clear: syntax data
- Clear: media curation
- Clear: alignment
- AI in the Bible Translation community

Logos Bible Software

- Faithlife: “We use technology to equip the Church to grow in the light of the Bible.”
- Logos: leading digital library for Bible study
 - 150k+ books
 - Desktop, browser, mobile
 - English, Spanish, Chinese, French, German, Korean, Portuguese



Content Innovation at Faithlife

- Linguistic annotation
- Knowledge management
- Data management
- Media, Atlas, more

Smarter, faster, cheaper ...



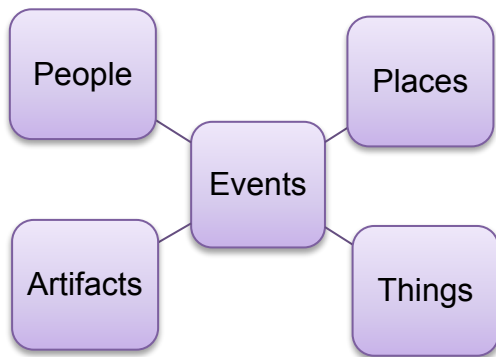
Timeline Events

Biographies

Saints

Church History

Historical Knowledge



Logos Controlled Vocabulary

Cultural Concepts

Preaching Themes

Concepts

Ancient Literature

Grammars

Systematic Theo.

Cross-References

Case Frames		Prop. Labels		Genre	Culture
Senses	Syntax	Discourse	Pericopes		
Bible text: words, lemmas, morph					
Aligned translations					
Biblical language & linguistics					

Bible Dictionaries

Commentaries

Lexicons

Reference Content

Biblical Places Data

- Unique identifier
- Preferred and alternate labels, localized
- Brief description
- Comprehensive biblical references
- Links to Bible dictionary content, Wikipedia
- Media
- Settings for biblical events
- Links to a world history timeline
- Geo-coordinates, nearby places, broader region

Challenges for Biblical Places Data

- Name variation
 - [Jerusalem](#) ≈ Ariel ≈ City of David ≈ City of God ≈ City of the Lord ≈ Hephzibah ...
 - Different names used in different eras
- Location
 - Some have moved over the millenia (modern vs ancient Ephesus)
 - Many are simply unknown
 - Region boundaries are approximate at best

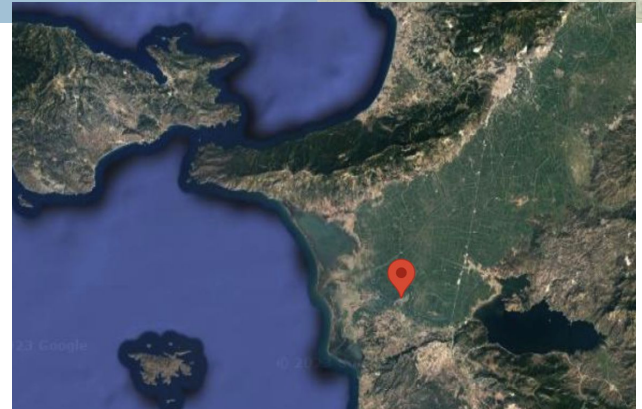
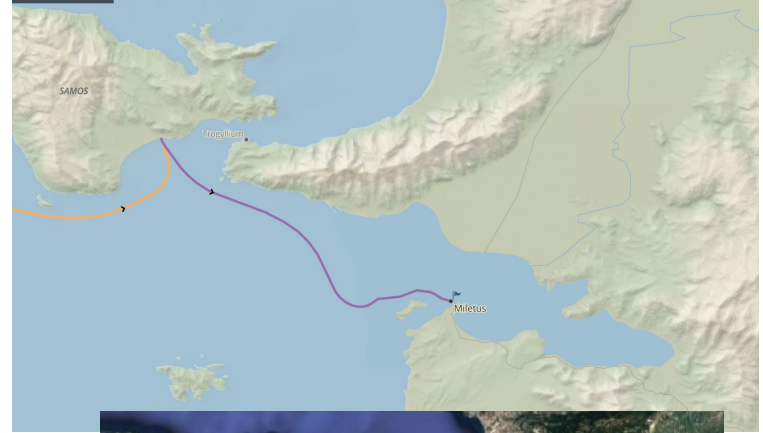
Logos Atlas

- Overview maps for 13 biblical eras
- Built around biblical events: “a map for every story that needs one”
- Metadata for passages, cities, events, hyperlinked to reference information.
- Era-specific road networks
- Distinguishes certain and uncertain location
- Zoomable, using [Mapbox](#)
 - Different city and region labels at different zoom levels
- Map designs by [Dr. Barry Beitzel](#)

Example: [the Battle of Elah](#)

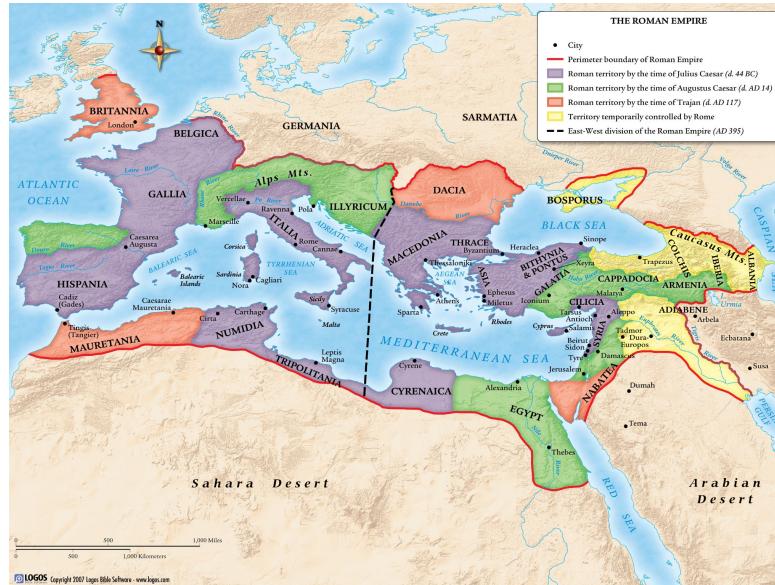
Challenges for Logos Atlas: Changing Geography

- [Miletus](#) is a *harbor* in Acts 20:15, but not any more.
- Modern satellite imagery is problematic therefore.



Other Geographic Features in Logos

- [Lexham Geographic Commentary series](#)
- Biblical Places maps



More Resources for Biblical Places and GIS

- [OpenBible.info](https://openbible.info) has a great collection of GIS data and an interesting Atlas project.

Clear Bible

About Clear Bible

- We work together with various partners to create specialized, user-driven, practical solutions to help translate the Bible faster and better.
- Technology consultants, innovators, engineers
- Biblical languages expertise, linguistic data, datasets, and more



Potential Applications for Greek & Hebrew Datasets



Standing on Others' Shoulders

Supported through
ETEN Innovation Lab



Frontend framework
builds on and adapts



Backend builds on and integrates datasets
from various partners, starting with:



Faith Comes By Hearing®



MACULA: Rich Linguistic Data

- **M**achine-**A**ssisted **C**orpus **U**nified **L**inguistic **A**nnotation
- Hebrew and Greek morphology, syntax, and semantics.
- Hebrew mapped to LXX Greek
- Reading environment incorporating interlinear display, search, and reference content

(Browser demo)

xml:id	ref	role	class	type	gloss	text	after	lemma	normalized	strong	morph	person	number	gender	case	tense	voice	mood	degree	domain	ln	fram
n40001001001	MAT 1:111		noun	common	[The] book	Βιβλος		βιβλος	Βιβλος	976	N-NSF		singular	feminine	nominative					33005	33.38	
n40001001002	MAT 1:112		noun	common	of [the] genealogy	γενέσεως		γένεσις	γενέσεως	1078	N-GSF		singular	feminine	genitive					010002 033003	10.24 33.19	
n40001001003	MAT 1:113		noun	proper	of Jesus	Ἰησοῦ		Ἰησοῦς	Ἰησοῦ	2424	N-GSM		singular	masculine	genitive					93001	93.169a	
n40001001004	MAT 1:114		noun	proper	Christ	Χριστοῦ		Χριστός	Χριστοῦ	5547	N-GSM		singular	masculine	genitive					93001	93.387	
n40001001005	MAT 1:115		noun	common	son	υἱοῦ		υἱός	υἱοῦ	5207	N-GSM		singular	masculine	genitive					10002	10.30	
n40001001006	MAT 1:116		noun	proper	of David	Δαυεῖδ		Δαυίδ	Δαυεῖδ	1138	N-PRI		singular	masculine	genitive					93001	93.91	
n40001001007	MAT 1:117		noun	common	son	υἱοῦ		υἱός	υἱοῦ	5207	N-GSM		singular	masculine	genitive					10002	10.30	
n40001001008	MAT 1:118		noun	proper	of Abraham	Ἀβραάμ		Ἀβραάμ	Ἀβραάμ	11	N-PRI		singular	masculine	genitive					93001	93.7	
n40001002001	MAT 1:211		noun	proper	Abraham	Ἀβραάμ		Ἀβραάμ	Ἀβραάμ	11	N-PRI		singular	masculine	nominative					93001	93.7	
n40001002002	MAT 1:212		verb		begat	ἐγέννησεν		γεννάω	ἐγέννησεν	1080	V-AAI-3S	third	singular			ao	active	indicative		23003	23.58	A0:n
n40001002003	MAT 1:213		det		-	τὸν		ὁ	τόν	3588	T-ASM		singular	masculine	accusative					92004	92.24	
n40001002004	MAT 1:214		noun	proper	Isaac	Ἰσαάκ		Ἰσαάκ	Ἰσαάκ	2464	N-PRI		singular	masculine	accusative					93001	93.180	
n40001002005	MAT 1:215		noun	proper	Isaac	Ἰσαάκ		Ἰσαάκ	Ἰσαάκ	2464	N-PRI		singular	masculine	nominative					93001	93.180	
n40001002006	MAT 1:216		conj		then	δὲ		δέ	δέ	1161	CONJ									89015	89.87	
n40001002007	MAT 1:217		verb		begat	ἐγέννησεν		γεννάω	ἐγέννησεν	1080	V-AAI-3S	third	singular			ao	active	indicative		23003	23.58	A0:n
n40001002008	MAT 1:218		det		-	τὸν		ὁ	τόν	3588	T-ASM		singular	masculine	accusative					92004	92.24	
n40001002009	MAT 1:219		noun	proper	Jacob	Ἰακώβ		Ἰακώβ	Ἰακώβ	2384	N-PRI		singular	masculine	accusative					93001	93.157a	
n40001002010	MAT 1:2110		noun	proper	Jacob	Ἰακώβ		Ἰακώβ	Ἰακώβ	2384	N-PRI		singular	masculine	nominative					93001	93.157a	
n40001002011	MAT 1:2111		conj		then	δὲ		δέ	δέ	1161	CONJ									89015	89.87	
n40001002012	MAT 1:2112		verb		begat	ἐγέννησεν		γεννάω	ἐγέννησεν	1080	V-AAI-3S	third	singular			ao	active	indicative		23003	23.58	A0:n
n40001002013	MAT 1:2113		det		-	τὸν		ὁ	τόν	3588	T-ASM		singular	masculine	accusative					92004	92.24	
n40001002014	MAT 1:2114		noun	proper	Judah	Ἰούδας		Ἰούδας	Ἰούδας	2455	N-ASM		singular	masculine	accusative					93001	93.173h	
n40001002015	MAT 1:2115		conj		and	καὶ		καί	καί	2532	CONJ									89017	89.92	
n40001002016	MAT 1:2116		det		the	τοῦς		ὁ	τούς	3588	T-APM		plural	masculine	accusative					92004	92.24	
n40001002017	MAT 1:2117		noun	common	brothers	ἀδελφοῦς		ἀδελφός	ἀδελφούς	80	N-APM		plural	masculine	accusative					10003	10.49	
n40001002018	MAT 1:2118		pron	personal	of him	αὐτοῦ		αὐτός	αὐτοῦ	846	P-GSM		singular	masculine	genitive					92004	92.11	
n40001003001	MAT 1:311		noun	proper	Judah	Ἰούδας		Ἰούδας	Ἰούδας	2455	N-NSM		singular	masculine	nominative					93001	93.173h	
n40001003002	MAT 1:312		conj		then	δὲ		δέ	δέ	1161	CONJ									89015	89.87	
n40001003003	MAT 1:313		verb		begat	ἐγέννησεν		γεννάω	ἐγέννησεν	1080	V-AAI-3S	third	singular			ao	active	indicative		23003	23.58	A0:n
n40001003004	MAT 1:314		det		-	τὸν		ὁ	τόν	3588	T-ASM		singular	masculine	accusative					92004	92.24	
n40001003005	MAT 1:315		noun	proper	Perez	Φάρες		Φάρης	Φάρης	5329	N-PRI		singular	masculine	accusative					93001	93.374	
n40001003006	MAT 1:316		conj		and	καὶ		καί	καί	2532	CONJ									89017	89.92	

Enhanced Resources++

Enhanced Resources++

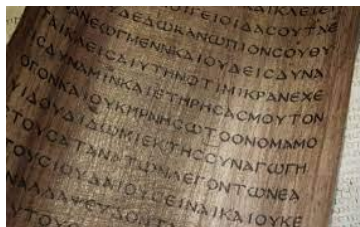
- Builds on the Enhanced Resources system in [SIL Paratext](#).
- Provide an open, extensible system for making additional resources available to translators and consultants.

מִצְנֶפֶת ^{12.1.1} n.f. turban—מִצְנֶפֶת—usu. of high priest; of king (Ezk 21:31), <OB> make Ex 28:4, 39] 39:28, לִבֵּשׁ dress with Si 45:12, place Ex 29:6]Lv 8:9, סֹר, hi. remove Ezk 21:31, <CSTR> מִצְנֶפֶת turban of linen Lv 16:4, שֵׁשׁ of fine linen Ex 28:39]39:28, <APP> כִּנְדֹר garment Ex 28:4, <PREP> בְּ of instrument, by (means of), with, + wind צֶנַח (the head) Lv 16:4; upon, + דָּרִיָּה be Ex 28:37, נָתַן give, i.e. place Ex 29:6]Lv 8:9(Sam) Ex 39:31, שֵׁם place Lv 8:9; פָּנֶיךָ אֵל at the front of, + דָּרִיָּה be Ex 28:37.

<COLL> מִצְנֶפֶת קִנְדֹרָה crown Ezk 21:31; Si 45:12, מִנְבִּיעָה headdress Ex 39:28 (+), מִכְנֶסֶת girdle Ex 28:4; 28:39]39:28 (+) Lv 16:4, אֶפֶד ephod Ex 28:4; חֲשֹׁן breastpiece Ex 28:4, מִעֵל robe Ex 28:4; Si 45:12, חֲנִיכַת nunic Ex 28:4; 28:39]39:28; Lv 16:4 (+), מִכְנֶסֶת breeches Ex 39:28; Lv 16:4 (both +). Also 4QapJoseph^h 12.
<SYN> מִצְנֶפֶת קִנְדֹרָה, מִכְנֶסֶת girdle, אֶפֶד ephod, חֲשֹׁן breastpiece, מִעֵל robe, חֲנִיכַת nunic.
- צֶנַח wind.



"give please" verb qal imp f sg + part interj נתן + נָא



בְּעֵמֶק V S T | ἐν κοιλάδι G (via בְּעֵמֶק) ÷ •
אַחֲרָיָה S T | ἀδεφου σου G (via אַחֲרִיךָ) | et post eum V •
מִנִּי V S (T) | ἐκ ἐμοῦ G | ἐκ χειρὸς G^{Mass} (assim-v 12) •

NOTES

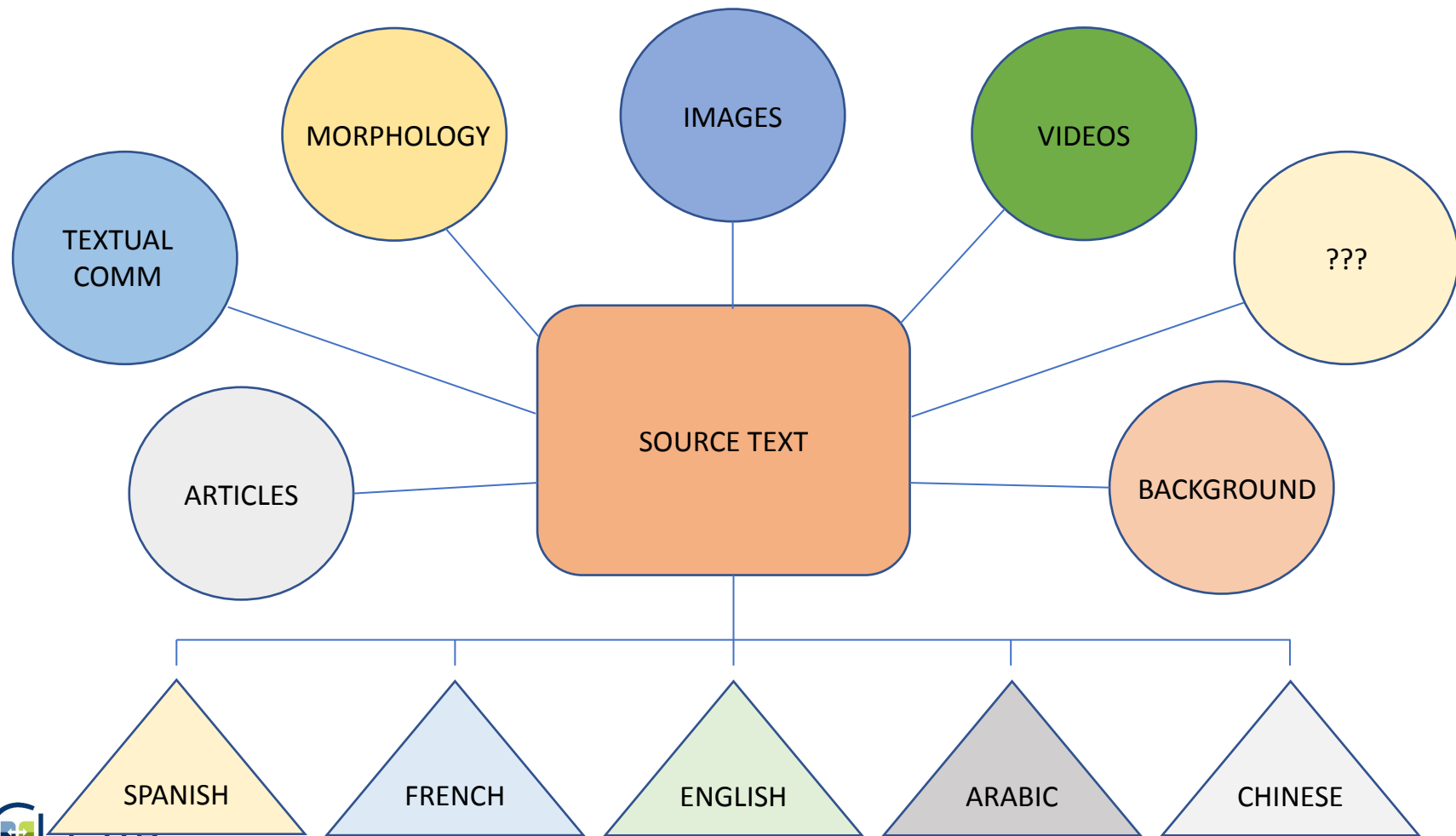
42. Psalms 42–43 form a single lyric consisting of three stanzas, each of which is followed by a refrain (42:5, 11, 43:5). Ps 42 may fairly be described as the biblical version of "The Dark Night of the Soul." # The psalmist, who had enjoyed God's consoling grace, finds himself in a state of extreme desolation because God has withdrawn his spiritual favors. He compares his wretchedness to the cheerless existence of Sheol.

2. As a hind. Reading (for MT *k'ayyal ta'arōg*, which is grammatically discordant) *ke'ayyelet' arōg*, and with Huesman in *Biblica* 37 (1956), 290, parsing *'arōg* as infinitive absolute functioning as a finite verb, as in Pss 17:5, 35:16, 49:4. This analysis is preferable to the supposition that a *t* was lost through haplography, as proposed by H. L. Ginsberg in BA 8 (1945), 56. Consult further Solá Solé, *L'infinitif sémitique*, § 58bis, pp. 95 f.

For the sequence *k ... kn*, cf. UT, 49:III:28–30, *klb arḥ l'gh klb tat timrh km lb 'nt atr b'l*, "Like the heart of a wild cow for her calf, like the heart of ewe for her lamb, so was the heart of Anath toward Baal."

Biblical *'ayyelet*, "hind," appears as Ugar. *aylt*.





Alignments in progress (UBS)

English	French	Spanish	Portuguese	Chinese	Russian
Swahili	Indonesian	Arabic	Tok Pisin	Tagalog	Cebuano
Amharic	Hausa	Hindi	Tamil	Telugu	Malayalam

Alignment Workflow

SIL Machine,
ClearEngine

ClearSuite,
ClearDashboard

To Be
Determined

Automated Alignment

Correction

Sharing

Parsing, segmentation,
morphological processing,
machine alignment.

Bulk correction by word.
Correction by verse. One or
more reviewers.

Sharing, discoverability, and
reporting.

Mass approve automatically aligned word pairs in ConcordanceTool

Fill in missed word pairs & fix wrong word pairs in AlignmentTool

ClearSuite

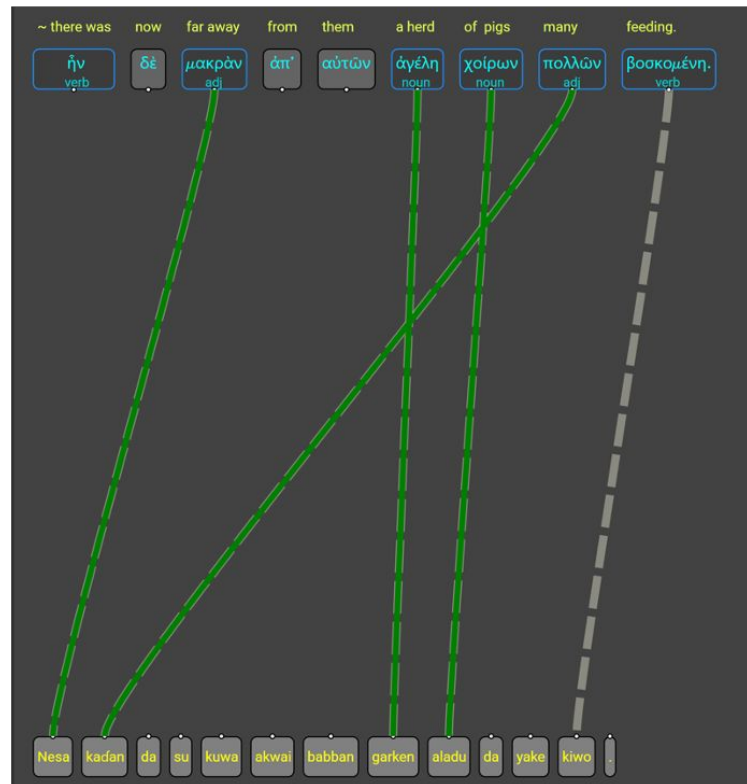
ConcordanceTool

Translation Words by Frequency

7329 - 3678 Occurrences

(7329 Occurrences) Ubangiji

3538 Occurrences	6626 Ubangiji	ḥḥḥḥ	Yahweh
2887 - 1813 Occurrences	526 Ubangiji	κύριος	of [the] lord
1369 Occurrences	127 Ubangiji	ḥḥḥḥ	Lord
1284 Occurrences	45 Ubangiji	ḥḥḥ	Yah
1087 - 1016 Occurrences	4 Ubangiji	ἐπιστάτης	Master
1011 - 968 Occurrences	1 Ubangiji	κυριεύω	he might be Lord.
837 - 834 Occurrences	(3678 Occurrences)	ce	
820 Occurrences	3542 ce	1_ ḥḥḥ	he said
797 Occurrences	62 ce	2_ ḥḥḥ	you say
795 - 793 Occurrences			



Known Needs for Alignment

- Standardized data format for sharing
 - Metadata (processing, provenance, etc...)
 - Alignment
- Centralized repository for discovering and using alignment data
- Community effort to collect and curate alignment data

Other Activities in the Bible Translation Community

- Bootstrapped draft translations and back-translations
- [AQuA](#) (Augmented Quality Assessment)
 - “... assists translators in reviewing local language drafts by objectively identifying potential issues and measuring a variety of relevant translation qualities.”

Connect

- Email: Sean.Boisen@Clear.bible
- LinkedIn: <https://www.linkedin.com/in/seanboisen/>
- Twitter: [@seanboisen](#), #BibleData
- Personal sites: [SeanBoisen.com](#), [SemanticBible.com](#) ([/Blogos](#))
- Clear Bible: <https://www.clear.bible>
- Git: <https://github.com/clear-bible>
 - Hebrew: <https://github.com/Clear-Bible/macula-hebrew>
 - Greek: <https://github.com/Clear-Bible/macula-greek>