A Computational, Historical-Critical Examination of ἐκκλησία in the New Testament

Matthew Munson
matthew.munson@uni-hamburg.de
ORCID: https://orcid.org/0000-0001-6211-2866

Ἐκκλησία in the NT: An historical-critical approach

- Jennifer Eyl, “Semantic Voids, New Testament Translation, and Anachronism: The Case of Paul’s Use of Ekklēsia”
  - ἐκκλησία does not yet mean “church” in the letters of Paul of Tarsus
  - Instead, she suggests the translation “assembly”
  - Her concern is that the later implications of the word “church” not be imported into the NT
  - Paul’s use is based in the Greek Old Testament (LXX) and Philo of Alexandria
  - Paul is interested to incorporate his new Christian communities into the existing Jewish ones
- The Money Quote: “If one were to believe the Liddell, one would think the entire semantic field suddenly changed in the middle of the first century...In all instances of ekklesia in Greek literature and the LXX, this word means *an assembly* (of some sort).” (334)

Investigate the claims

- Is ἐκκλησία translated as “church” in the NT?
  - YES! 108/114 times in the New Revised Standard Version
  - Always “church” when it clearly refers to a Christian group
- Do the lexica assert a unique translation for the New Testament?
  - YES!
  - **General**: LSJ, Middle Liddel; **NT**: BDAG, Louw-Nida
  - Other scholars lean towards church, even when they recognize a relationship to Jewish texts
- Did the semantic field of ἐκκλησία change “in the middle of the first century”?
  - Eyl says “No”
  - NT translations, Lexica (and most other scholars) say “Yes”
- What can we learn from a computational-linguistic approach?

Ἐκκλησία: A computational-semantic approach

- Distributional Semantics
- Firth, “You shall know a word by the company it keeps!”
- Harris, “The most precise way of determining a word’s meaning is by investigating the meanings of the words that occur along with that word.”
- Weaver, “But if one lengthens the slit in the opaque mask, until one can see not only the central word in question, but also say N words on either side, then if N is large enough, one can unambiguously decide the meaning of the central word.”
- Meaning from context

This Talk

- First half will discuss the computational method that I used
  - How to count
  - How to calculate significant co-occurrence
  - How to calculate word similarity
- The second half will focus on interpreting the results for ἐκκλησία
  - Close reading comparison of the results for the NT, LXX, and Philo

How to count? I.e., what is a word's context?

- I chose to use a fixed window N left and N right of the target word
  - Follows Weaver and typical procedure
  - Could have chosen other variable length logical units, e.g., sentence, paragraph, chapter, etc.
- But how large?
  - Studies have concluded from 1-500 words!
- And what type?
  - All words in the window are equally important? = Unweighted window
  - Words that are closer to the target word are more important? = Weighted window

Unweighted vs. Weighted: an example

Table 1.5: Unweighted context window

Count	1	1	1	1	0	1	1	1	1	0	0	0
Word	The	small	quick	red	fox	jumped	nimbly	over	the	lazy	brown	dog

Table 1.6: Weighted context window

Count	1	2	3	4	0	4	3	2	1	0	0	0
Word	The	small	quick	red	fox	jumped	nimbly	over	the	lazy	brown	dog

So let’s count!

Target Word	#1 Co-occurrent	#2 Co-occurrent	#3 Co-occurrent
ποιέω	ὁ	καί	αὐτός
ἔρχομαι	ὁ	καί	αὐτός
ἵνα	ὁ	καί	αὐτός
γίνομαι	ὁ	καί	αὐτός
ἔχω	ὁ	καί	αὐτός
κύριος	ὁ	καί	αὐτός
Ἰησοῦς	ὁ	καί	αὐτός
πᾶς	ὁ	καί	αὐτός
θεός	ὁ	καί	αὐτός
λέγω	ὁ	καί	αὐτός

How to Proceed? - Significant Co-occurrence

- Jacqueline Léon, “Significant collocation [=co-occurrence] is regular collocation between two items, such that they co-occur more often than their respective frequencies.”
- In the New Testament in general
  - ὁ (“the”) occurs once every 7 words
  - κύριος (“Lord”) occurs once every 200 words
- But in the neighborhood of the word ᾽Ιησοῦς (“Jesus”)
  - ὁ occurs once every 6 words (16% more often than one would expect)
  - κύριος occurs once every 10 words (2000% more often than one would expect)
- Distributional-semantic conclusion: the word “lord” tells us more about the meaning of the word “Jesus” than the word “the” does

Table 1.3: Log-likelihood ratio with a unweighted context window of 12L-12R.

	ποιέω	ἔρχομαι	ἵνα	γίνομαι	ἔχω	κύριος	Ἰησοῦς	πᾶς	θεός	λέγω
ποιέω	111.90	0.50	12.39	15.42	2.21	0.01	1.88	7.43	2.73	21.66
ἔρχομαι	0.50	28.51	11.16	0.16	2.22	0.78	14.18	16.46	46.89	14.69
ἵνα	12.39	11.16	16.07	4.32	21.77	2.99	0.84	10.30	1.71	15.28
γίνομαι	15.42	0.16	4.32	11.99	20.41	1.20	9.28	1.28	2.74	7.37
ἔχω	2.21	2.22	21.76	20.41	140.32	8.39	4.53	5.45	5.94	0.35
κύριος	0.01	0.78	2.99	1.20	8.39	102.47	99.90	5.39	0.26	33.40
Ἰησοῦς	1.88	14.18	0.84	9.28	4.53	99.90	0.68	5.99	15.96	143.06
πᾶς	7.43	16.46	10.30	1.28	5.45	5.39	5.99	226.53	9.99	80.87
θεός	2.73	46.89	1.71	2.74	5.94	0.26	15.96	9.99	142.80	49.82
λέγω	21.66	14.69	15.28	7.37	0.35	33.40	143.06	80.87	49.82	91.19

Target Word	#1 Co-occurrent	#2 Co-occurrent	#3 Co-occurrent
ποιέω	ποιέω	δένδρον	καρπός
ἔρχομαι	πρός	θεός	ὅταν
ἵνα	μή	ἐγώ	καί
γίνομαι	σεισμός	οἷος	ὀψία
ἔχω	ἔχω	χρεία	ἐξουσία
κύριος	ἐγώ	κύριος	ἰησοῦς
Ἰησοῦς	χριστός	λέγω	ἀποκρίνομαι
πᾶς	πᾶς	ὑποτάσσω	λέγω
θεός	θεός	ὁ	χριστός
λέγω	ἀποκρίνομαι	τίς	αὐτός

What now?

- Zellig Harris, “If we consider words or morphemes A and B to be more different in meaning than A and C, then we will often find that the distributions of A and B are more different than the distributions of A and C. In other words, difference of meaning correlates with difference of distribution.”
- So to find semantic relationship we need to find similarity of distribution
- The vectors of values that we saw previously are distributions
- So we need to find the similarity of these vectors

Cosine Similarity

$$CosSim(A, B) = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}}$$

	ποιέω	ἔρχομαι	ἵνα	γίνομαι	ἔχω	κύριος	Ἰησοῦς	πᾶς	θεός	λέγω
ποιέω	111.90	0.50	12.39	15.42	2.21	0.01	1.88	7.43	2.73	21.66
ἔρχομαι	0.50	28.51	11.16	0.16	2.22	0.78	14.18	16.46	46.89	14.69

CosSim(ποιέω, ἔρχομαι) ≈ 0.6430

Distributional Semantics as Historical Criticism

- Historical Criticism says that meaning is found primarily in historical context
- The ONLY “meaning” that distributional semantics returns is a word’s relationships with other words
- The only modern biases that are present are those that were introduced by the editor
- As you will see in a moment, the results still require a significant amount of interpretation
- But this is what we do as humanists!

Some important questions to consider

- We now have a process to calculate semantic similarity
- But what are the best parameters?
  - Context window size
  - Context window type
  - Significance algorithm
- To test this, we need to know what our results should look like, i.e., which words are similar to which other words
- In other words, we need a gold standard

Louw-Nida: the gold standard

- Johannes P. Louw and Eugene A. Nida, *Greek English Lexicon of the New Testament: based on Semantic Domains*
- Organizes all lemmata in the New Testament into semantic categories, e.g.
- “Supernatural Beings and Powers”
  - θεός (God)
  - ἄγγελος (angel)
  - σατανᾶς (Satan)
- When the clusters of similar words returned by these methods most resemble the Louw-Nida semantic domains, we have found our parameters!

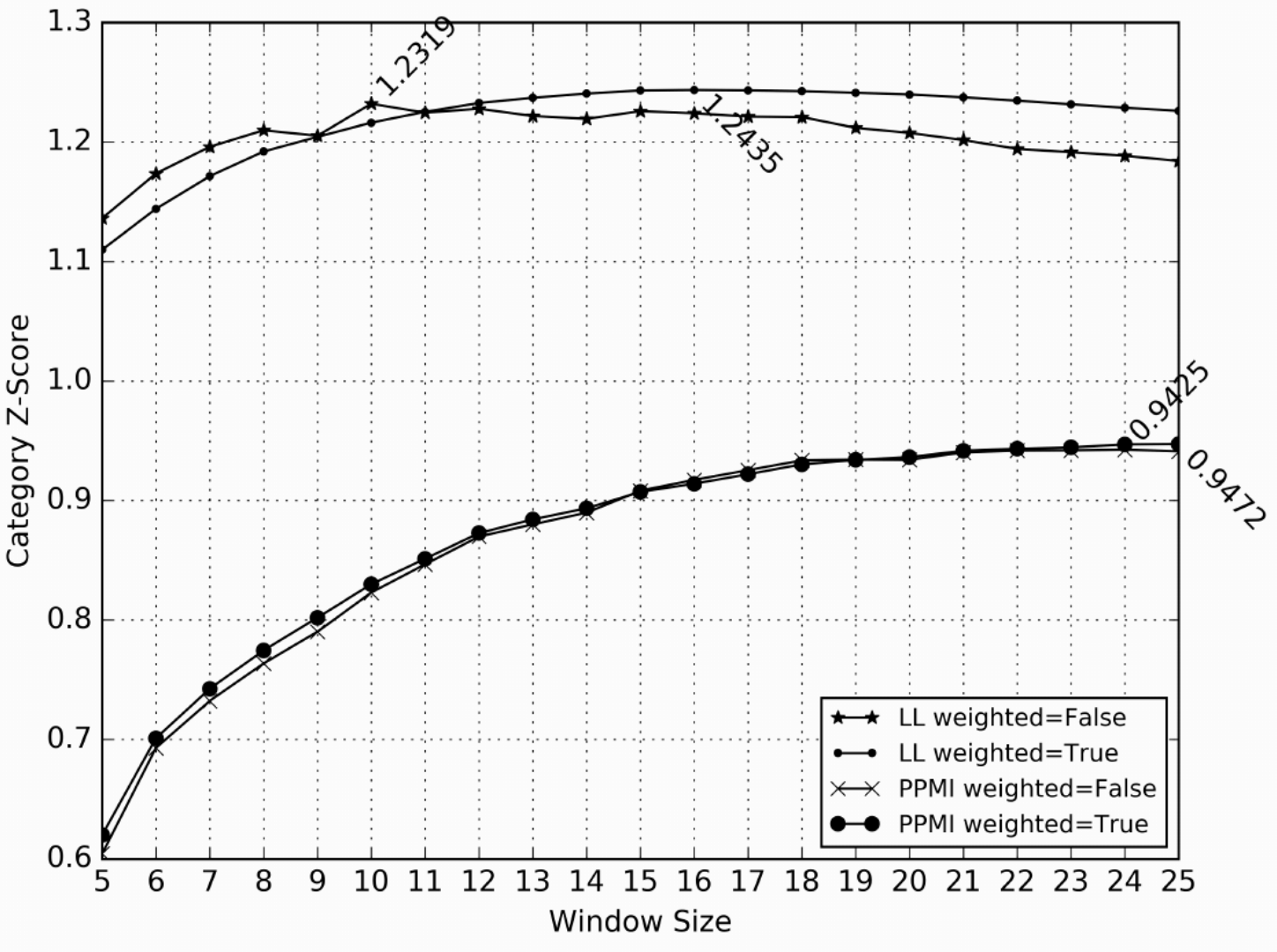

1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
Category Z-Score
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
Window Size
1.2319
1.2435
0.9425
0.9472
LL weighted=False
LL weighted=True
PPMI weighted=False
PPMI weighted=True

Table 1.11: Optimal window sizes for the different corpora

New Testament				Septuagint				Josephus		Philo		Plutarch		Perseus	
Lemmatized		Inflected		Lemmatized		Inflected									
10	1.07855	14	1.2404	4	2.0135	11	1.8052	33	1.8636	24	3.1348	47	2.1088	49	2.0639
11	1.07663	15	1.2430	5	1.9711	12	1.8063	34	1.8673	25	3.1356	48	2.1094	50	2.0646
12	1.08218	16	1.2435	6	2.1045	13	1.8064	35	1.8679	26	3.1479	49	2.1117	51	2.0650
13	1.07585	17	1.2432	7	2.0383	14	1.8059	36	1.8624	27	3.1418	50	2.1115	52	2.0648
14	1.07281	18	1.2424	8	2.0998	15	1.8045	37	1.8638	28	3.1460	51	2.1085	53	2.0647

Now, finally, we can investigate our problem

(the translation of ἐκκλησία)

Methodological Notes

- I lemmatized ἐκκλησία
  - I.e., I took every inflected form of ἐκκλησία and transformed it into the dictionary form
  - This results in a single semantic vector instead of several
- I compared similarity vectors
  - I.e., I calculated which words were the most similar to ἐκκλησία in each corpus
  - Then I compared the similarity vector for ἐκκλησία in one corpus with that in another
  - So I am comparing the relational meaning of ἐκκλησία in the different corpora
- This does not return synonyms
  - Instead, word similarity suggests that they share a common topic, e.g., “heavenly beings”
- Scholarly focus was on similarities between NT and LXX and/or Philo, so that will be my focus here

Table 1: 50 most similar words between the NT and LXX

ἀμπελῶνα	vineyard	ἀμπελῶνι	vineyard	ἐπεθύμησα	to covet	γινώσκεις	to know	ἀληθιναὶ	truthful
αἰγύπτιοι	Egyptian	ἐφάγομεν	to eat	εἶπω	to say	εὐθὺς	straight	ὄλεθρος	ruin
ἀφιέναι	to forgive	μαστιγώσουσιν	to whip	τέκτονος	craftsman	ἀθετεῖ	to refuse	ὅλος	entire
ψευδεῖς	to lie	ἔπαισεν	to strike	μερίδα	portion	ποιοῦντες	to do	θεοὺς	gods
ἀληθινόν	truthful	ἀφορῶντες	to look toward	καιρός	proper time	δείξω	to show	θέλει	to will
καθαρὸν	clean	ποταμῷ	river	ἀνεμνήσθη	to remind	ἔχομεν	to have	δαιμόνια	demon
ὀφείλει	to owe	κωφὸν	blunt	κληρονομήσω	to inherit	διακοσίων	200	κρεῖσσον	better
παρθένον	virgin	διανοήματα	notion	μένομεν	to stay	ὤμοσεν	to swear	προγόνων	ancestor
μόχθον	toil	πεποιθέναι	to persuade	ἔχουσα	to have	ἔμενεν	to stay	πόδα	foot
ἕτεροι	other	ἀνάπαυσιν	rest	ποίει	to do	ἔγνω	to know	λάβετε	to take

Now we need close reading!

- Black = Exodus
  - “Egyptian”, “toil”, “to whip”, “to strike”, “river”
- Medium gray = 10 commandments
  - “To lie”, “other...gods”, “to covet”, “to take (the Lord’s name in vain)”, “to do/make (idols, mercy, work)”
- Dark gray = promised land
  - “To inherit”, “to swear...ancestors”
- Here ἐκκλησία shares a semantic domain with words describing the events of the Exodus, i.e., Israel’s founding story

Table 1: 50 most similar words between the NT and Philo

μαθητὴς	pupil	εἶδον	to see	ἀφορῶντες	to look towards	λαλεῖν	to talk	ὅλῃ	complete
ἔτει	year	ψεῦδος	lie	κακὸν	bad	ὄχλον	crowd	ἔτος	year
υἱὸν	son	πολλαῖς	many	πρᾶξιν	deed	ἔπεσαν	to fall	ἱκανὸν	sufficient
ἐλθὼν	to come	εἴδει	form	ἄνθρωπος	person	καλῶς	rightly	γῆς	earth
ἤκουσας	to hear	ἱκανοῖς	sufficient	ἁγνὴν	pure	ἔκρινεν	to choose	καιροῖς	proper time
ἦλθον	to come	καθαρὸν	clean	ὁδῷ	road	καιρῷ	proper time	ὥρας	hour
ἐξῆλθεν	to come out	ἀληθῶς	truly	ἔθεντο	to place	ἄνθρωπον	person	δοὺς	to give
παρελθεῖν	to pass	ἀδελφοὺς	sibling	καλὸς	beautiful	εἶδεν	to see	εὐχαριστεῖν	to give thanks
χεῖρα	hand	ψευδεῖς	to lie	ἄδηλα	unseen	σαρκὸς	flesh	γεννᾶται	to beget
μικροῦ	small	ἑορτῆς	festival	θεὸν	God	ἴδωσιν	to see	καίσαρος	emperor

NT and Philo - communal religious celebration

- Medium-dark gray = time
  - year (ἔτει and ἔτος), “season, hour”, “proper time” (καιρῷ, καιροῖς)
- Dark gray = festivals
  - “festival”, “crowd”, “whole (community?, burnt offering?)”, “to give thanks”
- So ἐκκλησία shares a semantic domain with words about time and about gathering for giving thanks
- This suggests a shared focus on the religious festivals of the community

What does it all mean?

- Evidence we have seen supports Eyl's arguments
  - ἐκκλησία does not seem to have a new "semantic field" in the NT
  - Exodus theme demonstrates a clear point of contact between the NT and the LXX
  - Philo and the NT connect in terms of religious festivals
  - So ἐκκλησία appears to follow a Jewish model, but was it to integrate or to replace?
- This evidence suggests (DOES NOT PROVE) that we should rather choose a translation that we would use for Philo or the LXX instead of "church"
  - Assembly?
  - Congregation?
- It does not tell us, however, which translation to choose

Read more at...

http://nbn-resolving.de/urn:nbn:de:bsz:15-qucosa2-169575

Thank You!

One method: Pointwise Mutual Information

$$\mathrm{PMI}(t, c) = \log_2 \left(\frac{P(t, c)}{P(t)P(c)} \right)$$

- P(t,c) = # of times **t** and **c** co-occur ÷total # of co-occurrences
- P(t) = # of times **t** co-occurs with all words ÷total # of co-occurrences
- P(c) = # of times **c** co-occurs with all words ÷total # of co-occurrences
- But there is another...

Dunning's Log-Likelihood Ratio (1)

Hypothesis 1: $P(w^2|w^1) = p = P(w^2|\neg w^1)$ #independence

Hypothesis 2: $P(w^2|w^1) = p_1 \neq p_2 = P(w^2|\neg w^1)$ #dependence

$$log\lambda = log\frac{L(H_1)}{L(H_2)} = log\frac{b(c_{12}, c_1, p)b(c_2 - c_{12}, N - c_1, p)}{b(c_{12}, c_1, p_1)b(c_2 - c_{12}, N - c_1, p_2)}$$

$$b(k, n, x) = \binom{n}{k} x^k (1 - x)^{(n-k)}$$

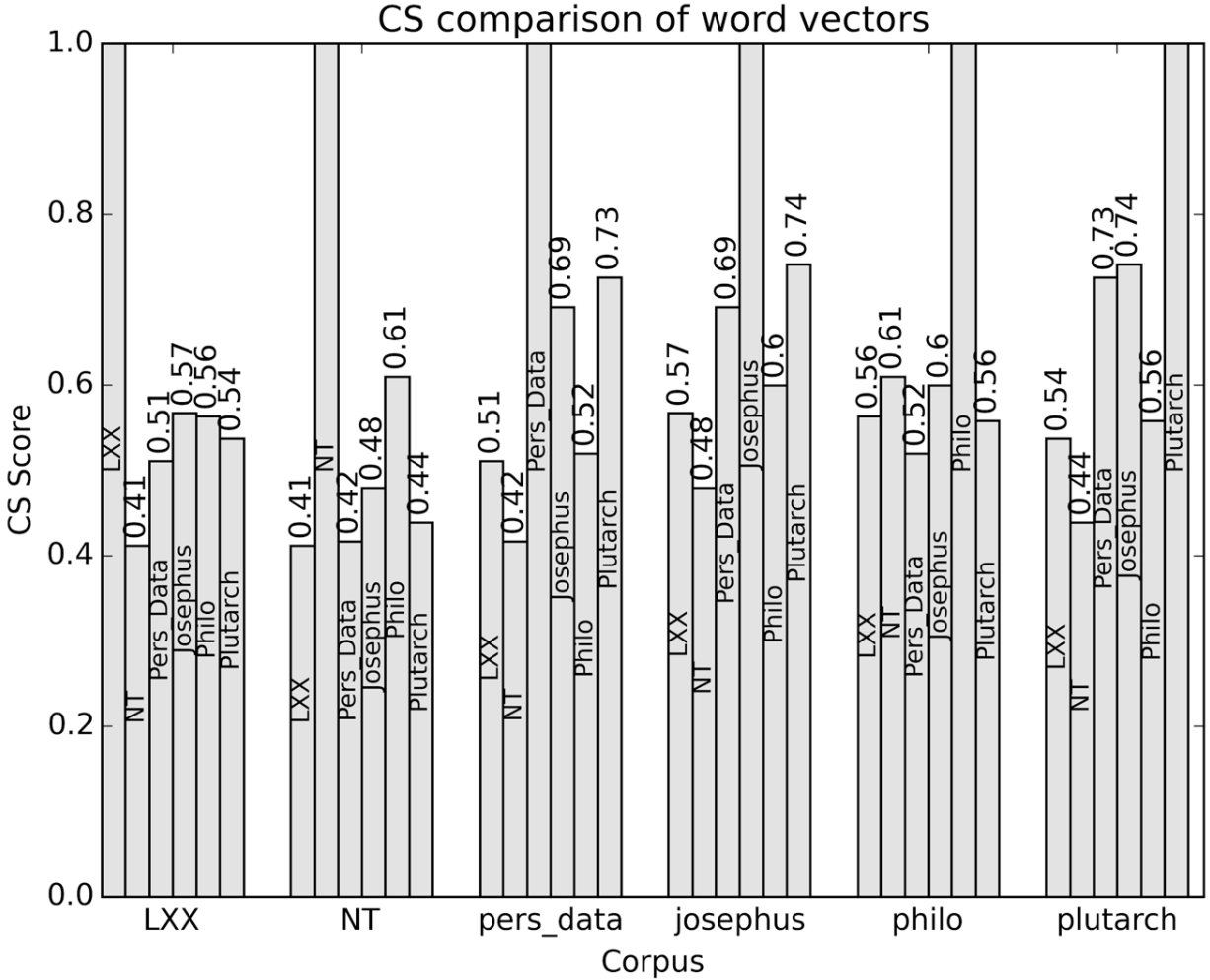
Dunning's Log-Likelihood Ratio (2)

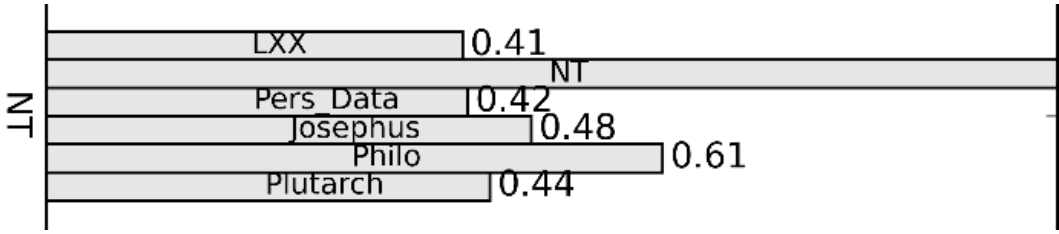
$$log\lambda = log\frac{L(H_1)}{L(H_2)} = log\frac{b(c_{12}, c_1, p)b(c_2 - c_{12}, N - c_1, p)}{b(c_{12}, c_1, p_1)b(c_2 - c_{12}, N - c_1, p_2)}$$

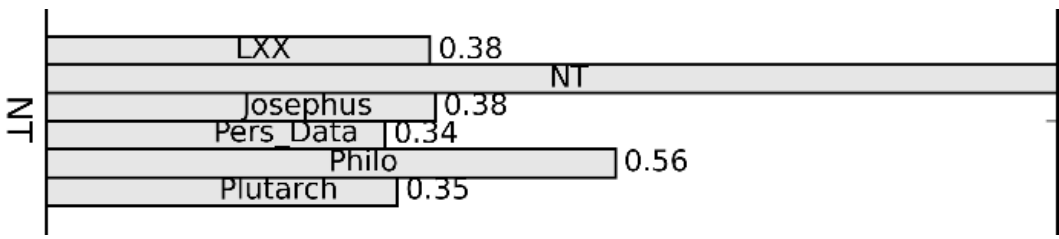
- c_1 = co-occurrences of target word
- c_2 = co-occurrences of co-occurrent
- c_{12} = co-occurrences of the two words
- N = total number of co-occurrences in the whole text
- $p = P(w^2|w^1) = P(w^2|\neg w^1) = c_2 \div N$
- $p_1 = P(w^2|w^1) = c_{12} \div c_1$
- $p_2 = P(w^2|\neg w^1) = (c_2 - c_{12}) \div (N - c_1)$
- b(k, n, x) has its maximum where $x = k \div n$, e.g., where $p = c_{12} \div c_1$

Louw-Nida: is this the best we can do?

- Lists only lemmata
  - So inflected forms will either have to be lemmatized or ignored
- Only for the New Testament
  - Will not mention a lot of words in other sources
  - However, their domain categorization depends on other sources besides the NT
- Their domain categorization depends on other sources besides the NT
  - Distributional methods depend only on the contexts that you feed them
  - If we use ONLY the New Testament, our distributional vectors and semantic similarities will depend ONLY on the NT
- But, at this point, it is the best we can do
- THE RESULTS!!!!


CS comparison of word vectors
CS Score
Corpus
1.0
0.8
0.6
0.4
0.2
0.0
LXX
NT
pers_data
josephus
philo
plutarch
LXX
NT
Pers_Data
Josephus
Philo
Plutarch
0.41
0.51
0.57
0.56
0.54
0.41
0.42
0.48
0.61
0.44
0.51
0.42
0.69
0.52
0.73
0.57
0.48
0.69
0.6
0.74
0.56
0.61
0.52
0.6
0.56
0.54
0.44
0.73
0.74
0.56


NT
LXX
0.41
NT
Pers_Data
0.42
Josephus
0.48
Philo
0.61
Plutarch
0.44


NT
LXX
0.38
NT
Josephus
0.38
Pers_Data
0.34
Philo
0.56
Plutarch
0.35