ANNALI
di storia dell'esegesi

39/1
2022

NETWORK SCIENCE
IN BIBLICAL STUDIES

ANNALI di STORIA dell'ESEGESI

Rivista semestrale
www.asejournal.net

Centro Italiano di Studi Superiori delle Religioni

Annali di Storia dell'Esegesi è classificata in fascia A per l'Area 11 e per l'Area 10 dall'ANVUR (Agenzia nazionale di valutazione del sistema universitario e della ricerca); e in categoria A e INT1 da ERIH (European Reference Index for the Humanities).

Tutti gli articoli sono sottoposti al processo di *peer-review*.

Gli articoli vanno inviati al seguente indirizzo: ase.rivista@gmail.com

Per le recensioni: ase.bookreview@gmail.com

Le riviste per scambio e i libri per recensione vanno inviati a: Redazione ASE c/o Cristiana Facchini, Dipartimento di Storia Culture Civiltà, Piazza San Giovanni in Monte 2, I-40124 Bologna.

Annali di Storia dell'Esegesi è indicizzato in: Année Philologique; ATLA Religion Database; Australasian Union List of Serials in Theological Collections; Index of Jewish Articles; Index Theologicus (IxTheo); Italian Bibliography of the History of Science; New Testament Abstracts; Old Testament Abstracts; Scopus.

Contents

Network Science in Biblical Studies

Abstract: This article introduces the special issue of ASE on network science and biblical studies. After a short presentation of network science and the concept of networks, the article discusses the application of network science in three domains: the natural and built environment of ancient Judaism and Christianity, the social networks of Jewish and Christian actors, and the analysis of textual corpora. In the next part, some technical terms of network science are clarified. The introductory article concludes with the presentation of the contributions to the special issue.

Keywords: Network Science, Social Network Analysis, Network Analysis of Texts, Historical Network Analysis, Vector Semantics

Abstract: Between the first and fourth centuries C.E. networks of like-minded individuals who shared certain religious beliefs and practices developed in both Jewish and Christian society. These networks emerged in the context of Roman imperialism which connected, integrated, and to some extent homogenized the various sectors of the population through road connections, infrastructure, and Romanization. Both rabbis and Christian leaders tried to distinguish themselves from certain aspects of their Graeco-Roman cultural environment by, at the same time, functioning within it and adopting some of its features. This study compares rabbinic and Christian leaders' networking activities within the shared Roman imperial context of the first centuries C.E.

Keywords: Rabbis, Bishops, Imperialism, Romanization, Communication, Travel

Abstract: Social network analysis is applied heuristically to understand how the Epistles of John attempt not only to promote their vision of the community's identity and theology, but

also to strengthen their own position in the Johannine social network and weaken the position of its competitors. Insistence on social distance from competing teachings is central to 1 and 2 John, while 3 John is occasioned by the Johannine School being marginalized by an assembly. The errands of these letters can be described in terms of network analysis as attempts to weaken certain social ties and strengthen others to allow the Johannine School more uncontested dispersion of information in the assemblies of the network.

Abstract: How successful was the Christian Church's mission to the Jews? The answer to this is a function of numerous factors, such as underlying demographics, fertility rates, competition, and (most importantly) when Christians and Jews severed their ties with one another. Some believe the split as early as 70 CE, but others present evidence that Jews and Christians regularly interacted until the 4th and possibly 5th centuries. In this paper, using computer network simulations that capture various factors, we run a series of models that estimate the proportion of Jewish converts to Christianity from approximately 40 to 600 CE. Our models suggest that whether one considers the Church's mission to the Jews a success largely depends on what one considers a "success." If one measures it solely in terms of raw numbers, then probably not. However, if one considers that the potential pool of Jewish converts was relatively small compared to Gentiles, it was perhaps more successful than many early Church scholars believe.

Abstract: The article joins the scholarly discussion about the meaning of righteousness language in biblical literature with consideration of changes in the concept from archaic Greek literature to fourth century Christian texts. The article seeks to showcase and evaluate how methods from the area of computational linguistics and distributional semantics can contribute to the discussion. The article suggests that, together with formal network models, namely word co-occurrence networks and similarity networks, the methods reveal changes in large corpora of textual data which are too subtle to be detected by close reading. On the other hand, some questions require or benefit greatly from combining distant and close reading methods.

Abstract: The verb "to circumcise" in Gen 17:24, and elsewhere in that chapter, appears in the niphal form, entailing a passive meaning, and the agent of the action remaining unknown. At the same time, Abraham's circumcision plays a central role in Judaism. It is not simply the prototype of all subsequent instances of a ritual that is central to Jewish identity; but it is also the starting point of a recursive chain of ceremonies by which any later circumcision acquires its religious significance. Hence, the importance of the agent of the first circumcision. First, I describe how this ritual is embedded in the network of concepts, entities, narratives,

precepts and further mental representations in Judaism. Subsequently, I shall present a selection of answers from various traditions to the question posed in the title. Then, I introduce a network of meanings, which shall serve as a linguistic model interpreting the agent-less niphal forms of the verb 'to circumcise' in Gen 17. This model yields different interpretations under various conditions, corresponding to various traditions. Consequently, it is argued to describe the computation taking place in the human mind, which is able to produce alternative interpretive traditions.

Abstract: This article puts forward the theory that the Bible is a network of cultural items (memes) that developed through an evolutionary process and has been transmitted for many centuries in a relatively stable form. Biblical cross-references reveal how the verses and passages of the Bible are connected into a network that is analogous to genetic networks in biology. The article presents a network model of cross-references, based on the *Treasury of Scripture Knowledge*. The article discusses the history of the database and the origin of the references; network statistics that yield insights into the history of the cross-references; node centrality statistics and the popularity of the respective verses in the history of Western biblical interpretation; and the modular structure of the network. The results are interpreted as tentative evidence for the collective behavior of populations in Western history being co-determined by a network of biblical memes.

Articles

Abstract: For the gospel passages too, as for any other historical document, it is impossible to interpret their meaning correctly without knowing their context, without therefore assigning them a precise place and time. The scholar of Christian origins is well aware that for the gospels the operation is considered to be mostly impossible, but he cannot resign himself to saying that the original context is irretrievably lost. There are, in fact, a few ways of attempting to recover it. One of these is to try to identify who the recipients of Jesus' words really were; to try in particular to understand to whom Jesus was addressing himself when he said "you". Some gospel pericopes allow this operation.

Abstract: New Testament scholars continue to identify the opponents of John of Patmos and his addressees at Smyrna and Philadelphia—despite John's testimony that "they say they are Judeans and are not, but are lying"—as a Jewish "Synagogue of Satan" (2:9, 3:9), not gentile imposters. Since 2001, David Frankfurter has identified the Smyrnean and Philadelphian opponents as "Pauline or neo-Pauline" gentiles who threatened the Jewish identity of John's con-

gregations. This proposal anticipated recent Pauline scholarship wary of Christian theological anachronisms, but the resulting "Paul within Judaism" shares more with Frankfurter's "John within Judaism" than with his "Pauline" foil. Refining Frankfurter's hypothesis to distinguish "neo-Pauline" inheritors from Paul promises a clearer view of John, his disciples, and his adversaries in relation to the contested reception of Paul's instructions for his gentile disciples.

Keywords: Synagogue of Satan, Revelation, Paul, Ephesians, Anti-Semitism

Abstract: The dreams of the patriarch Joseph (Gen 37), the dreams of the butler and the baker (Gen 40) and the dreams of Pharaoh (Gen 41) are significant dream examples in the Joseph-narrative. The Jewish exegesis of these dreams by Philo of Alexandria and the Early Christian exegesis of these dreams by Ambrose of Milan (337-397) will be discussed here. Next, we want to investigate whether Philo's exegesis of the dreams influenced Ambrose's exegesis or whether this is unlikely. Although we know that Philo influenced Ambrose far and wide, his influence on the interpretation of the dreams in the Joseph story seems to have been slight. Further investigation reveals that Ambrose relied on Origen for his exegesis of these dreams. As for the scholarly investigation of the latter, a complicating factor is that Origen's exegesis of Genesis is only partially left.

Keywords: Jewish and Christian Exegesis, Joseph's Dreams, Philo of Alexandria, Origen, Ambrose of Milan, Exegetical Influence

Book Discussion

Vojtěch Kaše, Nina Nikki, Tomáš Glomb

Righteousness in Early Christian Literature: Distant Reading and Textual Networks*

I. Introduction

The study of righteousness terminology from the perspective of Christian theology has generally taken the apostle Paul's thinking as a focal point. Paul is also at the center of this article although we will look at changes in righteousness terminology from a wider perspective. A traditional, nowadays outdated view held that Judaism was a religion of "works righteousness" which Paul replaced with a merciful God who either declared or made the undeserving people righteous through the atoning work of Christ. In this view, righteousness was understood in a legal and normative sense. Since the 1970s, especially through the influence of E. P. Sanders,[1] the so-called New Perspective on Paul has stressed that righteousness was, both in Judaism and in early Christianity, a *relational* term denoting membership in God's covenant. This perspective has emphasized the origin of the relational meaning of the δικ (*dik*)-word group in the Hebrew צדק (*ṣdq*) rather than in the extra-biblical Greek, with the latter still considered normative and judicial. Recently, the strong dichotomy between the relational and legal interpretations has been questioned, for example, by Michael Seifrid and Charles Lee Irons.[2] Other debates concerning biblical righteousness language have centered on the *salvific vs. punitive* understanding of the

---

[1] Sanders 1977.
[2] Seifrid 2001; 2004; Irons 2015.

term, that is, whether God's righteousness denotes the actions by which he redeems his people (due to his covenant faithfulness) or the punishments he metes out for wrongdoers. The debate concerning *declarative/imputed vs. infused* righteousness has a long history as well, although modern scholarship seems quite unanimous in locating this dichotomy in the later debates between Protestantism and Catholicism.

This article aims to contribute to the debates concerning righteousness terminology in early Christianity by demonstrating the potential of the distant reading approach to the topic.[3] The crucial part of this demonstration is providing a comparison of the results obtained by these quantitative methods from the area of computational linguistics and distributional semantics to those obtained by close reading, i.e., an established method of reading and interpreting individual texts.[4] In addition to this methodological validation, we discuss the potential of distant reading to reveal trends in large corpora of textual data that can be too subtle to be detected by close reading as well as the limits of this quantitative methodology. The article employs a comprehensive corpus of publicly available ancient Greek texts from the Homeric period to Late Antiquity which includes archaic and classical Greek, Jewish, Pauline, and other early Christian literature.

In the following, we first introduce the datasets and outline the methods employed in the quantitative part of the article. Subsequently, we discuss righteousness terminology in pre-Christian Greek sources, combining insights obtained by close and distant reading. An overview of righteousness terminology in the Hebrew Bible is also provided for contextualization, although the quantitative analysis is implemented solely on Greek texts. After this, we look at righteousness terminology in the Septuagint, which is considered to have an important role in mediating between the Hebrew Bible and Paul. Further, we look at Paul's use of the key term and provide examples of how textual network analysis can contribute to the main scholarly debates on the matter. Finally, we contextualize Paul's usage in comparison with other early Christian and Jewish authors of the Roman period.

II. Data and Methods

1. Datasets

The most important dataset for this article is a corpus of digitalized and publicly available ancient Greek texts named LAGT (Lemmatized

---

[3] The term distant reading was originally coined by Franco Moretti (Moretti 2013). Here we employ it in a broad sense to cover any computational text analysis method, especially when dealing with large portions of textual data (cf. Franzini *et al.* 2015; Underwood 2017).

[4] For a comparable approach see Špiclová, Kaše, 2020.

Ancient Greek Texts).[5] LAGT combines together textual data from two open-access resources.[6] In total, it consists of 1,457 works, 2,891,346 sentences and 31,248,866 words.

Each work within the corpus is associated with particular metadata (e.g., work title, author name etc.). Most importantly, each work has been assigned a unique CTS URN identifier,[7] which has allowed us to enrich the dataset by some additional metadata like cultural provenience (Christian, Jewish, or Pagan) or date. The date is typically expressed by means of one or two centuries (e.g., 4 c. BCE or 2/3 c. CE).

Some of the works in the corpus are either not suitable for our task at hand, because we miss important metadata for them (especially dating), or they are not relevant with respect to our research questions, because they originate from a very late period. Therefore, we mainly rely on a subselection from the corpus, consisting of texts which are dated within the range from 8 c. BCE to 4 c. CE (1,255 documents, 1,783,275 sentences, 21,086,074 words).

For most of our analysis, individual works have been grouped by authors. In some cases, we group together texts produced by more than one author, but which are sensible to treat together. For instance, we group together "Johannine literature", taking together the Gospel of John and the three New Testament epistles associated with this name (1-3 John).[8] Finally, we also group together the whole Septuagint or all minor New Testament texts which are not of crucial interest to us.

In addition to LAGT, we also work with PHI, a comprehensive dataset of ancient Greek inscriptions.[9] In total, it contains 218,162 inscriptions, 309,734 sentences, and 4,419,743 words. For 131,904 of the inscriptions, we were able to extract some sort of numerical dates, which allows us to explore some temporal trends in the data. To combat the substantial extent of uncertainty in dating (i.e., an inscription is commonly dated to a range covering one or even more centuries), we use a Monte Carlo model, which is an iterative mathematical tool for exploring probability distributions. It allows us to compare our results across 100 possible versions of the dataset, each containing a varying set of possible dates of origins for the inscriptions estimated based on their date ranges.[10] This approach allows us to analyze to-

---

[5] Kaše 2021a.

[6] In particular: Cerrato *et al*. 2020; Crane 2020.

[7] Blackwell, Smith 2019.

[8] Cf. Painter 2008.

[9] https://inscriptions.packhum.org.

[10] For details, see Kaše 2021b. For instance, in one simulation variant, an inscription dated to a Hellenistic period (i.e., an inscription which could be produced at any time between 330 BCE and 30 BCE.) will be dated to 3rd c. BCE, while in another simulation variant the same inscription will be treated as being produced during the 2nd c. BCE. See https://github.com/sdam-au/PHI_ETL.

gether both rather broadly and precisely dated inscriptions. Having these simulation variants, we can then ask whether a temporal trend of our interest is present across all of them or not. Following this, we group inscriptions within each variant by century and inspect what trends are robust across the simulation variants. In Figure 2-(b), we visualize results for each simulated variant of the dataset as one line. Places where the curve representing one word is thicker represent periods with lower extent of temporal uncertainty and therefore enable more inference.

2. Data Preprocessing

Before we delve deeper into these datasets, it is useful to explain some basic terms and methods on which the following analyses rely. First, to produce valuable results, distant reading methods require the textual data to be preprocessed in a certain way. Usually, the textual data have to go through tokenization, part-of-speech analysis (or: pos-tagging), and lemmatization.

Tokenization refers to a procedure of dividing the raw text into its constitutive parts. In our case, it takes place on two levels. First, the whole work is split into sentences. Subsequently, each sentence is divided into individual words, i.e., tokens. Pos-tagging is a set of algorithms assigning to each word its part-of-speech (POS) tag, i.e., whether a word in a sentence is a substantive (SUBS), adjective (ADJ), verb (VERB), punctuation (PUNCT), etc. Pos-tagging algorithms usually work in a probabilistic manner, assigning a tag for a word on the basis of the context of the sentence in which the word appears. These algorithms usually rely on some sort of machine learning, in which the algorithms are first trained on manually encoded data. Lemmatization is a process during which a word is coupled with its base (dictionary-like) form: e.g., Χριστοῦ is coupled with Χριστός (*Khristos,* "Christ"), ἐστιν with εἰμί (*eimi,* "to be"), etc. In most cases, this transformation can be done rather straightforwardly by means of a word-lemma dictionary. But there is also a significant number of instances in which this process is more complicated. For instance, a word found in a text might be either a substantive or a verb and therefore refer to two different lemmata. To deal with this issue, our lemmatization algorithm assigns lemmata on the basis of the pos-tags.

As a result of preprocessing, we obtain the textual data in a form in which each work is represented as a list of sentences and each sentence as a list of tokens. A token consists of the word itself, its pos-tag, and its lemma. Finally, we filter for words tagged as substantives, proper names, adjectives, and verbs, and extract their lemmata. We choose these types of words since we consider them to be loaded with the highest semantic value. This sort of filtering represents an alternative to

stop words filtering.[11] As an outcome of this extraction, for each work in the dataset, we finally obtain two versions of the textual data to be employed in all subsequent analyses: (a) lemmatized sentences: a list of sentences, with each sentence representing a list of prefiltered lemmata; (b) lemmata: a flat list of prefiltered lemmata, i.e., the same as (a), but without sentence divisions.[12]

Before we conclude this section, we have to be aware that because of the fact that the pos-tags and the lemmata are assigned automatically in a probabilistic manner and because the input textual data is highly diverse in genre, dialect etc., the performance of the pos-tagging and lemmatization algorithms is far from being perfect. Consequently, the usefulness of the resulting data always depends on the scale of the task at hand: to analyze a particular sentence from a work improperly tagged might cause serious problems, but once we make an analysis comparing large subcorpora, these minor issues should not skew the overall results.

In the following few sections, we will introduce a series of methods which will be employed below in the analytic part of the article. First, we will describe some methods based on word counts and frequencies. Special attention will be paid to the TFIDF algorithm, which enables to analyze word frequencies within one part of a corpus in respect to the corpus as a whole. Subsequently, we will shortly introduce a set of models approaching language data as vectors and matrices which can be analyzed by means of standard operations from linear algebra. This approach is especially valuable in allowing us to capture different types of relationships between words on a semantic level. Finally, we will introduce word-occurrence and similarity networks, which combine the above mentioned approaches with the methods of formal network analysis and visualization.

On a technical level, all distant reading analyses introduced in this article were implemented using Python 3 programming language.[13] Most of the scripts have a form of jupyter notebooks, consisting of independently executable cells, which facilitates easier code reusability.[14] All of the files are now publicly available online inviting future reuse. The scripts used to preprocess the data are part of the respective data repositories mentioned above.[15] The scripts containing the analyses introduced below are included in a repository accompanying this

---

[11] Stop words filtering is a process of elimination of specific words from quantitative textual analysis. Usually, stop words are the most common words in a language without any significant semantic load such as prepositions, conjunctions etc.

[12] All these preprocessing procedures were done by means of a set of scripts accessible in the LAGT repository on GitHub: https://github.com/sdam-au/LAGT.

[13] Van Rossum, Drake 2009.

[14] Kluyver 2016.

[15] See esp. Kaše 2021a.

article.[16] In the repository, there are also some supplementary figures and tables.

3. Word Counts, Frequencies, and TFIDF

Almost every distant reading analysis has to start with counting individual words. Comparing word counts for several words over the whole corpus can bring interesting insights (cf. Table 3 below). However, once we want to compare how certain words are used in different parts of the corpus, raw counts become insufficient, since the count depends on the size of these parts. This limitation might be overcome by *term frequency* (TF), in which the word count is divided by the total number of words. Therefore, when comparing usage of certain words across different parts of the corpus (for instance, comparing differences in usage between individual authors or from one century to another), we usually employ term frequencies, at least as a first step.

Looking at words with the highest TF, we usually realize that they tend to include generally widespread but semantically uninformative words, like εἰμί ("to be"). In order to overcome this limitation in part, we can employ TFIDF. In this algorithm, TF of a term within a specific part of a corpus is weighted by its distribution over the corpus as whole. This weighting is done by multiplication of TF by inverse document frequency (IDF). Document frequency is calculated as a number of documents containing the term. This value is then inverse and logarithmically normalized.

$$TFIDF = TFlog_2(IDF)$$

$TF_2(IDF)$ is a very efficient measurement for capturing terms highly specific for a certain part of a corpus or text. Below we repeatedly use this measurement to look at a certain number of words with the highest TFIDF values within preselected textual contexts of our interest.

4. Word Co-Occurrences and Vectors

According to distributional semantics, the meaning of a word is defined by its context of usage.[17] A context of usage may be operationalized in several different ways. One approach is to approximate it by means of sentences containing the word. Following this, in the subsequent analyses, we repeatedly focus on words co-occurring together

[16] Kaše, Nikki, Glomb 2021.
[17] See Lenci 2018, cf. Harris 1954.

within sentences. An alternative is to focus on n-grams or concordance windows of certain width surrounding the word.[18]

This simple logic stands behind a whole research area of vector semantics.[19] Perhaps the most important language model using vectors is word-document matrix, in which each row represents a word occurring within a document collection (i.e., its vocabulary) and each column represents a document within the collection. In many common applications, documents are equal to sentences within the collection. Individual cells then express the number of times a word appears within a document, i.e., a sentence. This matrix representation can be easily transformed into a second important vector model, word-word co-occurrence matrix, capturing how often any two words appear together across all sentences. These two matrix representations form the background for a broad variety of methods from the area of distributional semantics, some of which will be introduced below.

Analyzing the data in terms of vectors allows us to employ a broad palette of mathematical techniques from linear algebra. Simply speaking, treating a word as a vector means that we assign to it a position within a multidimensional space. This technique is known as word embedding and there are currently many approaches on how to generate such embeddings.[20] In its simplest form, a vector corresponds to row values within a word-document matrix, where each document corresponds to one dimension in this multidimensional space. It is then supposed that words with a similar position within this multidimensional space are also semantically closely related. A simple example might help to make this better understandable. Let us suppose that our document collection consists of the first two verses of the Gospel of John in English translation (NRSV):

> In the beginning was the Word, and the Word was with God, and the Word was God. He was in the beginning with God (John 1:1-2).[21]

In this case, the document-term matrix will consist of two columns, determined by the number of documents in the collection. The number of rows will be determined by the size of the vocabulary we choose to work with. In our example, we can decide that our vocabulary will consist only of two words: "Word" and "God". The word "Word" ap-

---

[18] N-gram is a sequence of a certain number (N) of co-occurring words in a text. A concordance window is a word context surrounding a selected word in a text. The width of the window defines how many words before or after the selected word are included.

[19] Jurafski, Martin 2017.

[20] E.g., Latent Semantic Analysis (LSA) (Deerwester *et al*. 1990; Landauer *et al*. 2007) and word2vec (Mikolov *et al*. 2013).

[21] Cf. Czachesz (2016), who uses the same two verses to explain word occurrence networks.

pears three times in the first document, but is completely missing in the second one. Therefore, the corresponding cells have the values 3 and 0. "God" appears two times in the first document and once in the second document. Therefore, the corresponding cells have the values 2 and 1 (see Table 1).

word	John 1:1	John 1:2
"Word"	3	0
"God"	2	1

Table 1. Word-document matrix example.

These cell values might be approached as defining direction and magnitude of two groups of vectors: (1) row vectors, representing each word in the vocabulary as a vector in space with number of dimensions equal to the number of documents; (2) column vectors, representing each document in the corpus as a vector in space with number of dimensions equal to the number of words in the vocabulary.[22] Since our corpus consists of two documents, the word vectors are defined only by means of two components, what makes possible to plot them directly within a 2-dimensional space, where the value for J 1:1 of given words defines its x axis vector component, while the value for J 1:2 of that word defines its y axis vector component. If you have a large number of words distributed across a large number of documents (sentences), this sort of approach might be applied to detect semantic similarities between words. By analogy, the same approach might be applied on the column vectors, which represent individual documents. In this case, the number of dimensions of the vectors corresponds to the number of words (i.e. number of rows of the matrix). This sort of method can be used to compare similarity between documents based on the vocabulary they employ. Figure 1 demonstrates both approaches. On Figure 1-(a) the vectors represent individual words and the documents define the axes, on Figure 1-(b) the vectors correspond to the documents and the words define the axes.

In the example above, the word-document matrix was based on word-counts. But it is important to realize that we can easily build a word-document matrix based on the TFIDF values. This approach is adopted below when we compare the usage of righteousness terminology across authors in a corpus of early Christian texts.

---

[22] In that respect, it is important to realize that, for example, the Pythagorean theorem, which might be used to calculate this magnitude) is equally applicable also for calculation of vector magnitude in a higher-dimensional space.
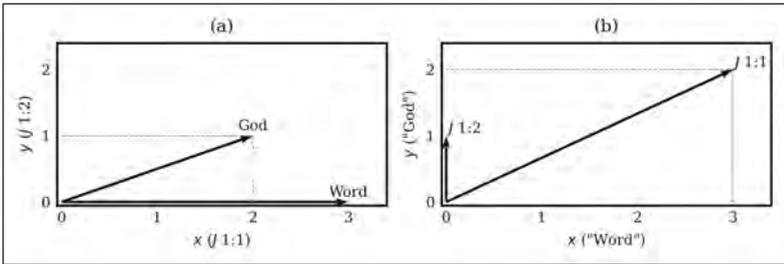
Figure 1. (a) words as vectors example; (b) documents as vectors example.

As already mentioned, the word-document matrix can be easily transformed into the so-called word-word co-occurrence matrix,[23] capturing how often any two words appear together across all documents within the collection. However, there are again a few different versions of the matrix. Strictly speaking, if we consider all their instances independently, "Word" and "God" co-occur 6-times in our document collection.[24] However, in some applications, it makes more sense to count their co-occurrence within a document only once. Further, we have to consider what values should constitute the diagonal of the matrix. Depending on the application, it might be either a) zeros, b) ones, c) number of documents within the collection containing the word or d) the total count of the words within the collection (see the results for all these four options in Table 2).[25]

	"Word"	"God"
"Word"	0 or 1 or 1 or 3	6 or 1
"God"	6 or 1	0 or 1 or 2 or 3

Table 2. Word-word co-occurrence matrix example.

5. Word Co-occurrence Networks and Similarity Networks

The methods outlined above can be fruitfully combined with formal network analysis methods. In this article, we mainly work with two types of networks: word co-occurrence networks and similarity networks. Word co-occurrence networks are networks generated from word

---

[23] In linear algebra terms, it is achieved by multiplication of the document-term matrix by its transpose (Strang 2016, 112-113).

[24] It is because in John 1:1, the word "Word" appears 3 times and the word "God" 2 times, and these 2 values are multiplied. In John 1:2, the two words do not appear together, so it does not contribute to the co-occurrence score.

[25] Notice that the resulting matrix is a symmetric square matrix with many important properties (see Strang 2016, 338-349).

co-occurrence data, e.g., in a form of a word co-occurrence matrix.[26] Once we obtain a word-word co-occurrence matrix capturing how often any two words appear together in a predefined context (e.g., within all sentences from one author within our corpus), we can use this data to generate a network in a very straightforward manner. This is because a word-word co-occurrence matrix can be treated as an adjacency matrix defining a network. An adjacency matrix is a square matrix with both columns and rows equal to nodes of a network.[27] A non-zero value within a cell indicates that there is an edge connecting the two nodes. In other words, whenever two words co-occur together within a sentence, there is an edge connecting the two words. Repeated co-occurrence of the same two words increases the weight of the edge.[28]

We also construct a similarity network of authors. Similarity networks might be defined as networks in which the presence and weight of an edge is based upon similarity of the two nodes under consideration. Such similarity can be based on many different things. In our case, we construct a similarity network of authors with edges connecting authors using the term δικαιοσύνη (*dikaiosynē,* "righteousness") in a similar way. This is based on a variant of a word-document matrix. The details of the procedure are explained below.

After this methodological section, we can turn to the analysis of righteousness terminology in the ancient Greek texts. We will proceed by combining insights based on the history of scholarship, insights achieved by our own close reading of the sources, and insights drawing on the quantitative methods sketched in this section. Initially, we will work mainly with word counts, word frequencies and TFIDF. Later on, we will also employ some vector and network based methods.

III. ANALYSIS OF RIGHTEOUSNESS TERMINOLOGY

1. Righteousness in Archaic and Classical Greek

The word δίκαιος (*dikaios,* "righteous") and related terms are etymologically connected with the verb δείκνυμι (*deiknymi,* "show", "indicate"). Based on this, David Hill suggests that the basic word in the group, δίκη (*dikē*), originally denoted "way indicated" or "that which is

---

[26] Mihalcea, Radev 2011.
[27] See Czachesz, Introduction, in this issue.
[28] We have also experimented with generating word co-occurrence networks from immediately neighboring words (i.e., a word adjacency network) (cf. Czachesz 2016). However, because of the limited space, here we focus on co-occurrence networks based on sentences.

customary".[29] The earliest appearance of the word refers to the goddess Δίκη (*Dikē*), daughter of Zeus (Hesiod, *Works and Days*, 256). According to Hill, Solon (ca. 630–560 BCE) was the first to "demythologize" the term. For him it meant "*a law*, the law of the universe, which is independent of man and cannot be evaded by him." After Homer, the term became increasingly connected to legal proceedings while also being used more widely in the sense of what is right or just.[30]

The LAGT corpus contains many other δικ-words. The most common of them are listed in Table 3. When we focus only on terms without the negative prefix ἀ-, the three most frequent terms are δίκαιος, δίκη, and δικαιοσύνη. In what follows, we will mainly work with these three terms.

lemma	count	translation
δίκαιος	12961	observant of custom and social rule, well-ordered, civilised
δίκη	6754	custom, usage
ἀδικέω	4745	to do wrong
δικαιοσύνη	4486	righteousness, justice
ἄδικος	2993	wrong-doing, unrighteous, unjust
δικαστής	2493	a judge
ἀδικία	2053	wrong-doing, injustice
δικαστήριον	1432	a court of justice
δικάζω	1054	to judge, to give judgment on
ἀδίκημα	1032	a wrong done, a wrong
δικαιόω	740	to set right
δικαίωμα	411	an act by which wrong is set right

Table 3. The most common δικ-words in AGT.

---

[29] Hill 1967, 98-99.
[30] Hill 1967, 99.

Hill distinguishes three stages in the use of the δικ-word group: the first word to appear was δίκη, then δίκαιος, and lastly δικαιοσύνη.[31] According to our temporal overview of these three terms in LAGT and PHI (see Figure 2, (a)-(b)), it seems that the first two terms started to be used extensively at the same time. However, it is interesting to notice that the rise in the usage of δικαιοσύνη from the 5th to the 4th c. BC can be documented both in literary texts (LAGT) and in inscriptions (PHI) in approximately the same time. This proves that the usage of the term δικαιοσύνη was not limited to literary works.
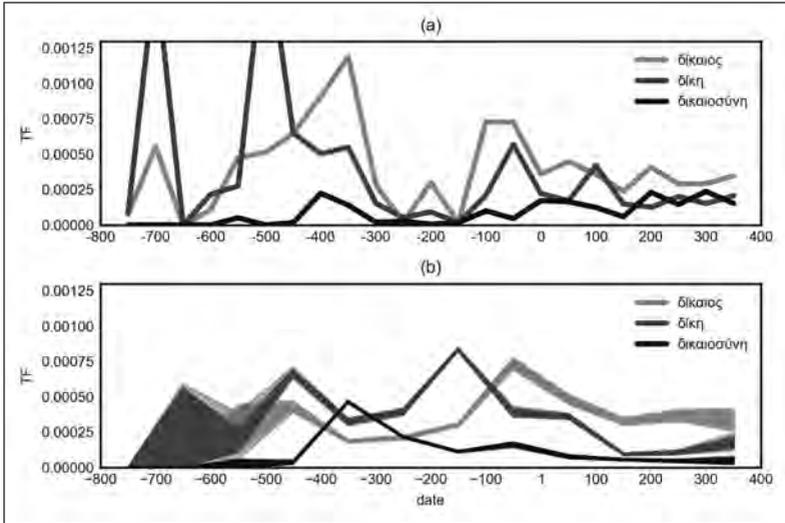
Figure 2. Term frequencies (TF) of δίκαιος, δίκη and δικαιοσύνη across corpora and subcorpora. (a) LAGT; (b) PHI. In (a), the peaks in the usage of δίκη around 700 and 600 BCE, depicted as exceeding the y axis range, are based on a very small amount of data and therefore should not be given too much emphasis

In the case of PHI (*Figure 2-(b)*), all simulated time series of the dataset are plotted together by means of a cumulative line plot. On the one hand, a high vertical dispersion of lines of the same color (e.g., lines plotting frequency of δίκη in the 7th and 6th c. BCE) indicates a high extent of temporal uncertainty in the data and therefore does not enable any meaningful inference concerning trends at the respective point in time. On the other hand, places where the lines of the same color tend to be close to each other indicate a small extent of temporal uncertainty

---

[31] Hill 1967, 100.

and therefore allow us to formulate robust inferences concerning trends in the data. Thus, we observe quite clearly that the usage of δικαιοσύνη increases extensively between the 5th and 4th century BCE. Taken together, this figure informs us that the term δικαιοσύνη was not unusually popular in Paul's time, at the beginning of the Common Era.

Charles Lee Irons emphasizes that the term δικαιοσύνη was used in many senses in extra-biblical Greek,[32] and that it included faithfulness as a subcategory of righteousness – making it unlikely that this was a purely Hebraic idea.[33] Herodotus (5th c. BCE) is the first to use the term consistently,[34] mentioning it eight times in five passages. He uses the term to denote, for example, honesty and justice in social situations and a personal quality (*Hist*. 6.86a, 7.164). It is also part of the collocation χράομαι δικαιοσύνη (*khraomai dikaiosynē*, "to practice righteousness") which is how Herodotus describes the custom of keeping political contracts (*Hist*. 2.151). The term is also used together with πιστότης to denote faithfulness and loyalty (*Hist*. 7.52). Irons, who questions the early use of δικαιοσύνη as a relational concept, classifies faithfulness as only one "species of righteousness".[35] The term is also used by Herodotus of "the quality of justice on the part of a judge" (*Hist*. 1.96) which Irons (following Havelock) believes to reflect the original judicial and royal context of the term.[36] Of other 5th century BCE authors, perhaps Antiphon[37] is most interesting as he actually ventures a definition of the term. For him δικαιοσύνη denotes observing the law of one's *polis*. According to Irons, this meaning of δικαιοσύνη as obedience to the law is common in extra-biblical Greek literature and is appropriated into Greco-Jewish literature, where the law is understood to be the Mosaic one.[38]

In the 4th century, δικαιοσύνη was used broadly by both philosophers and rhetoricians, which makes it possible and meaningful to

---

[32] The term is used variously to denote distributive righteousness, a general virtue, and of the observance of the law and as the goal of the law, and as human virtue coming from gods and rewarded by gods (Irons 2015, 105-107).

[33] Irons 2015, 106. This is in line with Irons's overall argument which is to dismantle the dichotomy between relational and normative uses of righteousness (Irons 2015, 84).

[34] In fact the earliest usage of the term in AGT is attested in *Aesopica*, conventionally dated to the 6th c. BCE.

[35] Irons's final aim is to demonstrate that Paul did not use the term δικαιοσύνη θεοῦ in a covenantal and relational sense and to enable a "renewed 'old perspective' understanding of Paul's 'righteousness/justification' terminology" (Irons 2015, 6). Irons understands righteousness in Paul as "the state of being legally recognized as dikaios before God, a state theoretically achievable by perfect good behavior and one which God requires of all humans if they are to avoid perishing and inherit eternal life." To remedy this and provide salvation, God gave Christ as "extraordinary righteousness" (Irons 2015, 7).

[36] For Herodotus, see Irons 2015, 87-89.

[37] Fragment 87.44 in Diels, Krantz *Die Fragmente der Vorsokratiker*. Note that this Antiphon is not necessarily the orator Antiphon.

[38] Irons 2015, 90-91.

employ some distant reading methods. Here we introduce a simple comparison of the usage by Isocrates and Aristotle. Table 4 contains 10 terms with the highest TFIDF values within sentences containing the term δικαιοσύνη in the works of these two authors. In both cases, we see there two basic types of words: (a) semantically crucial words from the domain of morality and virtues (δίκαιος, ἀρετή, ἀδικία, ἀνδρεία, σωφροσύνη, ἀγαθός, εὐσέβεια, σοφία);[39] (b) semantically less informative words like εἰμί, οἷος, ἄλλος, ἕτερος, πολύς, ἔχω, μέγας, οὗτος.[40] The proportion of terms in the first category covered here is striking. In both authors, the term with the second highest TFIDF within the sentences containing the term δικαιοσύνη is ἀρετή (*aretē*, 'goodness'), which demonstrates that both authors use the term in the context of virtues. Furthermore, both Aristotle and Iscorates tend to use δικαιοσύνη together with σωφροσύνη (*sōfrosynē*, 'prudence', 'discretion'). Finally, in the case of Isocrates, the algorithm captures a connection of righteousness with εὐσέβεια (*eusebeia*, 'piety'), while in the case of Aristotle we see an association with ἀνδρεία (*andreia*, 'manliness', 'bravery'). It is also remarkable that in the case of Aristotle, δικαιοσύνη often appears together with other δικ-words.

author	10 terms with highest TFIDF
Aristotle	δίκαιος, ἀρετή, ἀδικία, ἀνδρεία, σωφροσύνη, ἀγαθός, εἰμί, οἷος, ἄλλος, ἕτερος
Isocrates	σωφροσύνη, ἀρετή, εὐσέβεια, εἰμί, ἀδικία, σοφία, πολύς, ἔχω, μέγας, οὗτος

Table 4. 10 terms with the highest TFIDF in sentences containing δικαιοσύνη in Aristotle and Isocrates.

In addition to these words, the table also contains a number of terms from the second category – words with a lower semantic load. We encounter such terms repeatedly throughout the rest of this article. Some of them could perhaps be treated as stop words and excluded from the analyses. However, we keep them in for several reasons: first, they might have a semantic value in specific contexts; second, a decision to exclude a word from an analysis makes any quantitative analysis prone to cherry-picking (i.e., an activity with an attempt to skew the analysis to obtain deserved results by modifying the dataset).

---

[39] *Dikaios*, "righteous, just"; *aretē*, "goodness"; *adikia*, "injustice"; *andreia*, "manliness"; *sōfrosynē*, "prudence"; *agathos*, "good"; *eusebeia*, "piety"; *sophia*, "wisdom".
[40] *Eimi*, "to be"; *hoios*, "such as"; *allos*, "other"; *heteros*, "other", *polys*, "many"; *ekhō*, "to have"; *megas*, "large"; *houtos*, "this".

The close association between righteousness and virtues in the works of Isocrates and Aristotle has also been observed by Irons. According to him, Aristotle makes a distinction between universal justice and particular justice. Universal justice denotes for him a superset of all virtue – as ἡ τελειά ἀρετή – (1129b30, 1130a9–10)[41] and is also defined as that which is lawful (1129a34).[42] Particular justice on the other hand takes place in human interactions and denotes proportionality (1131b17–18), equality, and fairness (*iustitia distributiva*). According to Irons, there is thus a relational aspect to righteousness as Aristotle understands it (esp. 1161b6–7).[43] In Isocrates, piety (εὐσέβεια) is directed towards gods whereas righteousness (δικαιοσύνη) belongs to the sphere of human interaction.[44] The distant reading method employed above (terms with the highest TFIDF values within sentences containing the key term δικαιοσύνη) is unable to capture these types of semantic differences. This does not, however, mean that it cannot be done by applying some other distant reading methods. However, this would require a different type of input data, that is, dependency treebanks, which are annotated, syntactic trees of sentences. Using such data, we could focus on the syntactic position of δικαιοσύνη with respect to other terms of interest. Despite substantial progress on this front over the recent years,[45] data annotated in this way are available only for a small subset of ancient Greek texts. Alternatively, we could focus on the usage of additional word types, such as prepositions, which have been excluded from the dataset as we employ it in this article.

2. Righteousness in the Hebrew Bible

Although the current article applies distant reading methods only to Greek terminology, it is necessary to shed some light on Hebraic righteousness language, as well. The use of the δικ-root by Paul and other early Christians is mediated through the Septuagint, which is (roughly) a Greek translation of the Hebrew Bible. Discussion of the Hebrew Bible word group rooted in צדק (*ṣdq,* translated with terms denoting righteousness)[46] has revolved around two major issues. First, there is

---

[41] For a similar understanding, see Plato, *Resp.* 4.433b.

[42] See also Hill 1967, 99-100.

[43] Irons 2015, 95.

[44] Irons 2015, 91. A good representative of the same type of thinking from a later period is the first century BCE historian Diodorus Siculus, who uses the term δικαιοσύνη in many contexts. It represents moral goodness in "its most generic sense" (*Hist*. 1.2.8) and appears in the conventional pair εὐσέβεια καὶ δικαιοσύνη which covers human relations to both gods and other humans (*Hist*. 1.2.2. etc.) (Irons 2015, 106).

[45] Cf. Celano 2019; Gorman 2020.

[46] The following meanings are offered by Brown-Driver-Briggs:

צדק *ṣadaq, qal*: "be just", "be righteous", "have a just cause", "be in the right", "be justified", *nif'al*: "be put right", "be justified", *pi'el*: "justify," "make to appear righteous" *hitpa'el*:

a seeming discrepancy between *salvific* and *punitive* uses of the term. Second, the term has been understood – like its Greek counterpart – to denote either "accordance with a norm" or "fidelity in a relationship".[47] Another discussion has revolved around the question of "ethical" vs. "forensic" uses, the latter meaning a declared or pronounced status of righteous and the former denoting enacted righteousness. As for the latter, it has been stressed that "forensic" interpretations are anachronistic, unrealistic, and reflective of later Catholic-Protestant debates.[48]

Mark Seifrid has criticized any attempt to find a general meaning, a *Grundbegriff*, for a particular word-root, which would apply to all usages (i.e., word-equivalence). Instead, he stresses that terms should be viewed in their contexts, which actually determine their meaning.[49] This emphasis resonates well with the approach of distributional semantics, which we employ in the distant reading parts of this article. Concerned by words in their context, Seifrid notes that צדק very rarely appears together with בְּרִית (*bᵉrît*, "covenant"), which severely problematizes the covenantal readings of the term.[50] Seifrid thus denies that righteousness language is merely a *Verhältnisbegriff* denoting *Gemeinschaftstreue*, that is, a relational concept signifying fidelity to the covenant.[51] Seifrid still admits the existence of a covenant connection, but one that is mediated through the language of love חֶסֶד (*ḥesed*) and loyalty אֱמוּן (*ʾemûn*). These terms appear more often in covenant context and fall under the overarching idea of righteousness. Thus "[a]ll 'covenant-keeping' is righteous behavior, but not all righteous behavior is 'covenant-keeping.'"[52]

"justify himself", *hifʿil*: "do justice", "declare righteous", "justify", "vindicate the cause of", "save", "make righteous", "turn to righteousness"

צֶדֶק *ṣedeq*: "rightness", "righteousness", "what is right, just, normal", "justness", "justice", "ethically right", "justification", "deliverance", "victory", "prosperity", "covenant-keeping"

צְדָקָה *ṣᵊdāqâ*: "righteousness", "vindication", "truthfulness", "ethically right", "justification", "salvation", "prosperity", "righteous acts"

צַדִּיק *ṣadîq*: "just", "righteous", "right", "correct", "lawful"

[47] Seifrid 2001, 415-416. For a covenant fidelity view, see Hill 1967, 93-98.

[48] Seifrid 2001, 422. Ziesler (1972, 22) thus: "When we speak of the forensic righteousness of an accused person, we mean he is innocent, not guilty… his rightness or righteousness is a matter of status in, or relationship to the court. When a judge, however, is righteous or just, the reference is not so much to his status as to the way in which he acts in court." See also Seifrid 2004, 66: "One simply does not find bare pronouncements in Hebrew Scriptures or in Paul. The biblical writers are interested in justice enacted."

[49] Seifrid 2001, 418, 422. (A position originally argued by Barr 1961).

[50] Out of the 283 appearances of בְּרִית and 524 appearances of צדק the two are linked together only seven times (Seifrid 2001, 423). The simple fact that Noah and Abraham were attributed righteousness before and outside the covenant is also telling (Seifrid 2001, 425). The Qumran literature, on the other hand, links righteousness and covenant, but stresses that righteous deeds in fact bring salvation (Seifrid 2001, 434-435). In rabbinic literature, obedience is also connected to covenant, while צְדָקָה begins to mean almsgiving (Seifrid 2001, 438-439).

[51] Seifrid 2004, 43.

[52] Seifrid 2001, 424.

Importantly, a relational and covenantal understanding of righteousness language has been used to emphasize the difference between the Greco-Roman and biblical uses of the term, respectively. Seifrid, like Irons above, emphasizes that such a dichotomy is not realistic – both Greco-Roman and biblical uses have normative and relational elements. There is a difference, however, in that God is portrayed as the source and executor of righteousness in biblical language. In a biblical understanding, righteousness cannot be conceived of or distributed without the involvement of God.[53]

Seifrid also notes that צדק terminology appears much more frequently with God's ruling and judging activity (שפט, *šṗṭ*), which supports, for example, Hans Heinrich Schmid's argument that righteousness language is primarily rooted in *creational theology*.[54] Seifrid links the different aspects of righteousness to a view of God as "divine king, who has determined to secure the good and beneficial order of creation", which explains the salvific overtones of the term as well as the joined judicial, legal, and executive acts of righteousness by God.[55] Seifrid stresses that righteousness stems from God's role and actions as creator, ruler, and judge of the entire earth – a perspective strengthened by the fact that gentiles, and not only Israelites, can also be connected to צדק: God's acts of righteousness "do not represent mere 'salvation' for Israel, or even merely 'salvation.'"[56]

As for punitive and salvific usages, the different forms of צדק tend to reflect the different usages. The masculine צֶדֶק (*sedeq)* denotes the abstract concept of "right order" or "that which is morally right", while the feminine צְדָקָה (*ṣəḏāqâ*) is used of concrete matters such as "righteous acts" or "vindicating judgements".[57] The punitive form appears, according to Seifrid, altogether 15 times in the Hebrew Bible, the majority of the occasions referring to God as a righteous (צַדִּיק *ṣadîq)* judge who punishes the wicked.[58] The feminine צְדָקָה (*ṣəḏāqâ*), on the other hand, appears in vindicative connections.[59] Seifrid, in fact, suggests that the emphasis on this feminine form as "actual righteousness" hindered older scholars from detecting the punitive connections.[60] Irons, however, emphasizes that God's enactment of justice has both punitive and salvific elements: "righteousness of God" refers to "God's justice

---

[53] Seifrid 2004, 43-44.

[54] Appearing together in 142 instances. Here Seifrid sides with Schmid's (1968) suggestion that righteousness is closely related to creational theology (Seifrid 2001, 425-426).

[55] These functions are joined in the ancient world in general (Seifrid 2001, 426-427).

[56] Seifrid 2001, 440-441; Seifrid 2004, 42-43.

[57] Seifrid 2001, 428.

[58] Especially noteworthy are the confessions of God's righteousness by wrongdoers after being punished (e.g., Exod 9:27).

[59] E.g., Judg 5:11, Deut 33:21, Ps 22:32, 31:2, 36:7, 36:11, 40:11, 51:16, 71:2, 71:15 etc.

[60] Seifrid 2001, 430.

in executing judgment on the enemies of his people and thereby vindicating his people in the face of their oppressors."[61]

3. Righteousness in the Septuagint

The Septuagint translates the Hebrew root צדק quite consistently with forms of δικαιοσύνη/δίκαιος: out of the 476 possible instances 462 are translated with the term.[62] Of the existing 139 appearances of the צדק root in the book of Psalms, for example, the Septuagint translates 136 with forms of δικαιοσύνη/δίκαιος and 3 with forms of ἐλεημοσύνη (eleēmosynē, "mercy").[63] (The δικ-root itself appears altogether 1791 times in the Septuagint).[64] This faithful equivalence has led some scholars to suggest that the δικ-root in the Septuagint receives its semantic meaning, not from extra-biblical Greek, but from the allegedly relational/covenantal Hebrew צדק.[65] David Hill, for example, suggests that the δίκαιος-words "underwent considerable expansion and change of meaning through being consistently used to render the Hebrew root צדק", and one of these changes was that the term acquired the meanings of mercy, loyalty, and faithfulness in covenantal contexts.[66] As became evident above, this interpretation of the Hebrew term has been strongly criticized lately, which puts the idea of δικαιοσύνη as a calque (a loan translation) in dubious light.[67] According to Irons, covenant faithfulness "is not one of the Hebraic meanings that has been introduced into the Greek δικ-group among Greek speaking Jews or Christians of the period under investigation."[68]

Seifrid lists some ways that the numerous translators of the Septuagint both retained and changed the meanings of the Hebrew Bible. Perhaps most noteworthy is his conclusion that the Septuagint often retains the judicial sense of the צדק and, especially, the idea of vindicative justification, which, according to him is a major difference between

---

[61] Irons 2015, 178.

[62] Hill 1967, 104.

[63] Palola 2007, 113. Hill (1967, 104) claims "the Greek term retained the flavour of 'strict justice' throughout its entire usage" which explains that certain instances of צְדָקָה were translated in the LXX with ἔλεος (eleos, "pity") and ἐλεημοσύνη (eleēmosynē, "mercy").

[64] Seifrid 2004, 39 (who does not specify which manuscript or critical edition is used). The δικ-root is sometimes also used to translate other Hebrew terms such as ישׁר (jšr, "straight forward", "upright") and, on nine occasions, חֶסֶד (ḥsd, "an act of kindness or loyalty between people" or "the steadfast loyal love of Yahweh for his people") (Hill 1967, 105-107).

[65] Ziesler 1972, 67.

[66] Hill 1967, 108-109. The other changes were that the δικ-root became to be used of the divine, whereas secular Greek made this connection only in very early times. Also, while both use the word group for conformity with a standard, in classical Greek this standard is social but in the Septuagint it is divine.

[67] Irons 2015, 74-75.

[68] Irons 2015, 83.

Hellenistic and biblical usage. As for changes, Seifrid notes that the book of Job offers an ethicizing translation and the Psalms of Solomon lay aside vindicative connotations stressing retributive language instead. Seifrid also stresses that צדק is never translated with σωτηρία (*sōtēria*, "salvation").[69]

It is fruitful to turn to the distant reading methods and to compare the usage of the term in the Septuagint with non-biblical Jewish authors writing in Greek, such as Philo and Josephus. As before, we can focus on 10 terms with the highest TFIDF values (see Table 5). It appears that the usage by Philo and Josephus is closer to that observed in the case of classical Greek authors like Isocrates or Aristotle than to the Septuagint. Similarly to these classical authors, there is again an explicit association with ἀρετή. In addition to that, Philo's list includes σωφροσύνη and ἀνδρεία, which he shares with Aristotle. To some extent, this matches the analysis by Seifrid, according to whom "a major shift away from the biblical perspective" took place with Josephus (as well as Philo), who uses righteousness language in the Greek sense and describes righteousness as a virtue to be learned and practiced.[70] Furthermore, it appears that in Josephus there is a close connection of righteousness with a religious domain, since the two terms with the highest TFIDF value are εὐσέβεια and θεός. On the surface, from these three subcorpora, the punitive element appears to be most clearly present in the Septuagint through an association of righteousness with κρίμα and κρίνω. Thus, it may challenge Seifrid's claim of the dominance of vindicative rather than punitive righteousness in the Septuagint. However, further measures are needed to determine the actual ratio between punitive and vindicative use here.[71]

author	10 terms with highest TFIDF
Septuagint	κρίμα, ποιέω, κύριος, ἀλήθεια, θεός, ἀδικία, ὁδός, κρίνω, ἀνομία, λαός
Philo	φρόνησις, ἀρετή, σωφροσύνη, ἀνδρεία, ἰσότης, ἀδικία, εἰμί, ὁσιότης, ἄλλος, ἀγαθός
Josephus	εὐσέβεια, θεός, εἰμί, ἀρετή, δόξα, πολύς, ἀμάνου, ἄλλος, οὗτος, ἀνήρ

Table 5. 10 terms with highest TFIDF in sentences containing δικαιοσύνη in the Septuagint, Philo, and Josephus.

---

[69] Seifrid 2004, 49-52.
[70] Seifrid 2004, 45-49. According to Seifrid, Philo and other Hellenistic Jewish authors yield similar results.
[71] See, for example, the vindicative use – with κρίμα and κρίνω – in Ps 7:9 and 72:1-2.

4. Righteousness in Paul

The meaning and role of righteousness in Paul's writings has been the source of vigorous scholarly debate for centuries. Paul has often also been the starting point for theories concerning righteousness in the Hebrew Bible and the Septuagint. Irons, for example, points out that the covenantal theories of righteousness began with Paul and spread from there to the Septuagint and Hebrew Bible.[72] The continuity or discontinuity between (and beyond) these writings is, indeed, a major question for any study on the matter.
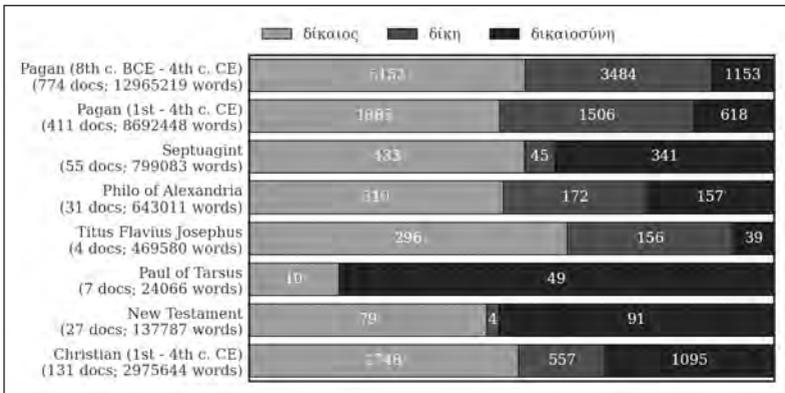
Figure 3. Proportional counts of δίκαιος, δίκη and δικαιοσύνη across LAGT subcorpora.

Looking at the proportions of the usage of δίκαιος, δίκη and δικαιοσύνη across various subcorpora (See Figure 3), we observe that Paul's preference for δικαιοσύνη is exceptional. It cannot be explained either by his relying on Jewish or "pagan" Greek sources. Paul's usage even skews the overall usage of these terms within the New Testament.

Before discussing the semantic meaning of δικαιοσύνη language in Paul, it is important to note that the relative centrality of the theme in Paul has been much debated. While later Western Protestantism insists on the centrality of the concept, this does not reflect Paul's message in its original setting. On the contrary, it has been argued since the time of Albert Schweitzer and William Wrede that Paul's "theology" centers on the idea of the believer's participation in Christ and that the concept

---

[72] Irons 2015, 36-39.

of righteousness is only a "subsidiary crater".[73] It has been suggested that righteousness is "a polemical doctrine" taken up by Paul mainly in debates with so-called Judaizers concerning the requirement of the Mosaic law on Gentiles.[74]

Philip Esler, for example, notes that righteousness terminology is completely missing from Paul's first letter, 1 Thessalonians, where Paul even replaces the concept with other terms when alluding to Scriptural passages.[75] This may mean that Paul approached people of gentile background with a different type of message, one overall less embedded in Scriptural content. Esler suggests that before Galatians and the debate described therein "[Paul] already had a well-developed theology in 1 Thess, and now [in Galatians] he seeks to re-work this in relation to the acquisition of righteousness. The reactive nature of his procedure is apparent."[76] Esler argues that the concept of righteousness was widely applied in Second Temple Judaism as a symbolic descriptor of the positive distinctiveness of Jewish identity in relation to Gentiles. Paul's aim in Galatians, in particular, is to appropriate this valuable term for his own group.[77] Esler thus stresses the fluidity of the term, which does not, however, nullify an attempt to construct some kind of general logic to Paul's use of the concept.

E.P. Sanders, a pioneer of the covenantal understanding of righteousness, sees Paul's use of the concept in terms of his model of *covenantal nomism,* where inclusion in God's people is "by grace", as it were, and law observance is necessary to remain in this relationship.[78] What separates Paul from Scriptural usage in Sanders's model is the way Paul places righteousness and Law observance on the "getting in" side of the covenant ("by the works of the law no one will become righteous", Gal 2:16), whereas for the Hebrew Bible and the Septuagint righteousness is covenant maintenance language ("staying in").[79] Partly at least, this can be viewed as Paul's polemical effort to portray his opponents and Judaism in general as promoting "works righteousness".[80] For Sanders, too, Paul's use of righteousness is reactive and not constitutive. Paul emphasizes the lack of righteousness in all people

---

[73] Schweitzer 1998 [orig. 1931], 225.

[74] Wrede 1907, 122. For discussion, see Nikki 2019, 185-186.

[75] In 1 Thess 5:8 Paul alludes to Isa 59:17, Wis 5:18, but replaces the term righteousness with faith and love (Esler 1998, 156-157).

[76] Esler 1998, 171.

[77] Esler 1998, 170-177. See also N. Nikki 2019, 161-162.

[78] Importantly, the covenant also includes atonement measures to be used in case of transgression. For a summary of covenantal nomism, see Sanders 1977, 422-423. The background of the discussion arises from the longstanding misrepresentation by Christians of Jews as promoting legalism and works-righteousness.

[79] Sanders 1977, 205, 624-625; 1983, 45.

[80] Nikki 2019, 182.

(Rom 1:18–3:20) as a result of his leading conviction that Christ is the savior of all humankind: it follows from this that every person must be in need of salvation.[81]

Seifrid criticizes Sanders' covenantal understanding of righteousness. He disagrees with Sanders's basic idea that all Jews were considered righteous on the basis of their membership and "normal" actions in the covenant, and that God rejected these "works of the Law", because he wanted to save Gentiles through faith in Christ. According to Seifrid, the boundary between righteous and unrighteous was not drawn between Jews and non-Jews, but often *within* the Jewish people, which creates room for a normative use of the term.[82] Most importantly, Seifrid questions the mutual exclusivity of covenantal nomism and judicial interpretations, stressing that Paul did not find all people sinful after his discovery of Christ as savior, but simultaneously with it. Overall, Seifrid argues that Paul follows biblical rather than extra-biblical Greek usage of the term. With his emphasis on vindicative righteousness (e.g., 1 Cor 6:11) Paul differs, for example, from Josephus' Hellenistic usage of the term.[83]

As we proceed to the application of distant reading to the research problem, it should be emphasized that when applying this sort of method to such a small and widely studied collection of documents as the letters of Paul, we can hardly expect the methods to produce substantially new insights. In this context, the application of these methods rather serves to validate them, enabling us to explore to what extent our methods are able to capture certain semantic features. As above, we begin by looking at the 10 terms with the highest TFIDF values within sentences containing the term δικαιοσύνη (Table 6). First, we observe here a connection with πίστις (*pistis*, "faith", "trustworthiness") and πιστεύω (*pisteuō*, "to believe", "trust"), which may refer to covenant fidelity.[84] A close association is also observable between δικαιοσύνη and νόμος, which can be interpreted either in covenantal or judicial vein. In this context we should also note the prominent position of ἁμαρτία (*hamartia*, "sin"), which is also open to interpretation.[85] We should hesitate to construct any strong conclusions based on these distant reading observations. We decided, for example, to focus on 10 terms for our comparison; had we decided to focus on 5 or 15 terms instead, we could have observed slightly different patterns.

---

[81] Sanders 1977, 443.

[82] According to Seifrid (2004, 64-65), only "figures of notable piety from the past" were considered righteous and when the "righteous" are contrasted with the "wicked" for example in Psalms, it is a matter of obedience to the law.

[83] Seifrid 2004, 52.

[84] For the changing meanings of faith terminology in Rom (which can at times even be viewed as parallel to δικαιοσύνη), see e.g., Sanders 1977, 490-491.

[85] For sin as disobedience in a covenant context, see Sanders 1977, 111-116.

author	10 terms with highest TFIDF
Paul of Tarsus	πίστις, ἁμαρτία, νόμος, λογίζομαι, θεός, βασιλεύω, δοῦλος, παρίστημι, πιστεύω, Χριστός

Table 6. 10 terms with the highest TFIDF in sentences with δικαιοσύνη in Paul.

To elaborate these observations further, we can now employ another distant reading method: word co-occurrence networks. As sketched in II. Data and Methods, we can easily generate a word co-occurrence network from a word-word co-occurrence matrix capturing co-occurrences of any two words within a list of sentences. This way we can create a word co-occurrence network for each author in our dataset. With such a network, we can explore the network position of the term δικαιοσύνη in respect to other terms.

Since the network consists of a substantial number of nodes, making it difficult to inspect it visually, we subsequently focus only on a small part of it, namely the immediate neighborhood of the term δικαιοσύνη in the network. In Figure 4, we see such a subnetwork for Paul's authentic letters. This subnetwork has been extracted in the following way: first, the algorithm extracted 5 nearest neighbors of the key term δικαιοσύνη (nearest neighbors are terms most commonly appearing together within sentences). Subsequently, for each of these 5 terms, this procedure was repeated, as we extracted their own 5 nearest neighbors. In principle, this procedure can generate 31 terms in total. However, in practice, this number is usually significantly lower, since there are overlaps between the terms.

We see that the nodes within the subnetwork are to a large extent the same as the ones we have among the 10 terms with the highest TFIDF (Table 6). In this case, however, the terms are accompanied with structural information. It has been suggested that the structural information captured by word co-occurrence networks might be used as a proxy for word association networks.[86] Thus, for instance, we observe that δικαιοσύνη is strongly associated with πίστις, ἁμαρτία, and νόμος. This combination is unsurprising in the light of Paul's typical argumentation concerning the role of faith to the exclusion of the law in attaining righteousness (e.g., Gal 3:21-22; Rom 3:21-22, 4:13, 9:30-31, 10:5-6).

---

[86] Galea, Bruza 2015. The nearest neighbors of a term within such a network might serve as an approximation for terms closely associated with the term in the mind of the writer. In effect, we can imagine that reading these texts can modify association networks in the minds of the readers (or listeners). Thus, word co-occurrence networks might be viewed as "models of the network of associations that the text generates in the readers' or listeners' minds" (Czachesz 2016, 45).
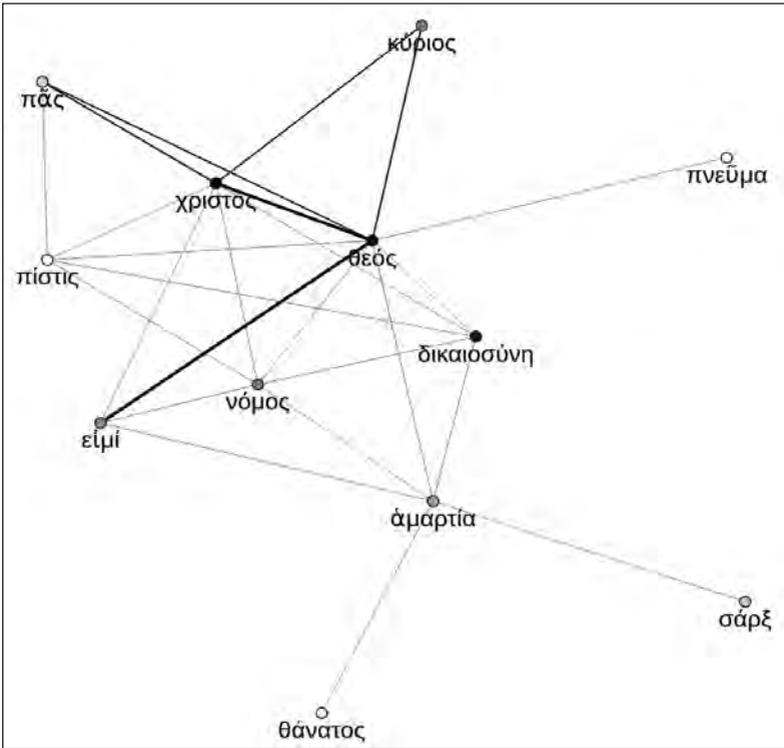
Figure 4. Subnetwork consisting of nearest neighbors surrounding the term δικαιοσύνη within the co-occurrence network based on the authentic letters of Paul (see the main text for a more detailed explanation).

But this is not all. We further see the nearest neighbors of these terms. In particular, we see that νόμος maintains a close association with πίστις. This is expected since they represent diametrically opposing paths to righteousness (see Phil 3:9). Another interesting aspect of this subnetwork is the usage of ἁμαρτία, which appears to be closely associated with νόμος and δικαιοσύνη, but not with πίστις. In addition to that, ἁμαρτία mediates a connection with θάνατος (*thanatos,* "death") and σάρξ (*sarx,* "flesh"), denoting mortality and corporality. The advantage of word co-occurrence networks is the ability to capture structural properties of the semantic relationships. However, it is more challenging to use this method to compare such semantic relationships across more authors, because of the limited comparability of the networks.[87]

---

[87] The limited comparability of word co-occurrence networks is caused by the fact that their size (number of nodes) and density (number of edges) is proportional to the size of the

To proceed further, we will introduce other distant reading methods, which will allow us to compare quantitatively Paul's usage of the term δικαιοσύνη with the usage of other authors. Again, we will draw on the TFIDF measurement employed above. But this time we will do so in a much more elaborated way, combining it with vector semantics. Drawing on the distributional hypothesis, we assume that when two authors use similar terms in the context of the term δικαιοσύνη, this implies that they also tend to understand the term δικαιοσύνη in a similar way.

As explained in section II, distributional semantics typically transforms textual data into their vector representations. Such vector representations can be mathematically analyzed and subjected to a comparison. Following this logic, a list of TFIDF values of terms appearing within sentences containing the term δικαιοσύνη within works from one author can be treated as a vector, defining one particular position of a point in a multidimensional space.[88] This allows us to compare TFIDF values of a list of terms across more than one author, with values of each author being represented by one point in a shared multidimensional space. This position is then subjected to a comparison. A standard measurement for this comparison is cosine similarity, which compares any two vectors by measuring the angle between them.[89] As an output of the cosine similarity measurement, we obtain a similarity matrix comparing any two authors under consideration. Our dataset as a whole consists of 269 authors, which ultimately corresponds to a similarity matrix with 269 columns and 269 rows. Since it is difficult to inspect such a matrix visually, here we focus only on its subselection consisting of authors or subcorpora we discussed in the previous sections: Isocrates, Aristotle, the Septuagint, Philo, Josephus, and Paul (Table 7).

The similarity values have to be always interpreted in relational terms, since the value depends on the size of the vocabulary and the extent to which this vocabulary is shared across authors. We should not be surprised by the relatively high similarity values connecting Isocrates and Aristotle. We could be perhaps more surprised by the extent of similarity between Philo on the one hand and Aristotle with Isocrates on the other (0.57362 and 0.63929 respectively). In the case of both Josephus and Philo, we observe a higher similarity with classical authors than with the Septuagint. Paul reveals the highest similarity with the

_____

input textual data. For instance, these differences disable application of standardized network measurements, like betweenness centrality of a certain node (e.g., the term δικαιοσύνη), across such networks, what would be otherwise very valuable information.

[88] The number of dimensions equals the number terms on the list.

[89] Cosine similarity relies on the dot product operator, which guarantees that the similarity value will be high only when the two vectors have large values in the same dimensions, and not when they share a lot of zero values. Cf. Jurafsky, Martin 2017, 279-281.

	Isocrates	Aristotle	Septuagint	Flavius Josephus	Philo Judaeus	Paul of Tarsus
Isocrates	1	0.50845	0.25758	0.49441	0.57362	0.15754
Aristotle	0.50845	1	0.26054	0.36762	0.63929	0.15322
Septua-gint	0.25758	0.26054	1	0.32254	0.30049	0.32178
Flavius Josephus	0.49441	0.36762	0.32254	1	0.43474	0.23478
Philo Judaeus	0.57362	0.63929	0.30049	0.43474	1	0.17135
Paul of Tarsus	0.15754	0.15322	0.32178	0.23478	0.17135	1

Table 7. Similarity matrix of vectors based on words with the highest TFIDF values in sentences containing δικαιοσύνη – a subselection of pre-Christian authors.

Septuagint (0.32178), not only within this subselection of authors, but among all pre-Christian extra-biblical texts in our dataset. However, the extent of this similarity (0.32178) is still lower than the one we observe between Philo and Josephus and the classical authors. On the one hand, the results affirm the distance between Paul and extra-biblical Greek authors, but on the other hand, it also sets Paul apart from biblical language. The complexity, even obscurity, of Paul's argumentation on righteousness has made scholars look to biblical texts for answers. Our results suggest that this should be done cautiously, since Paul also diverts from previous tradition.

5. Righteousness in Early Christianity

Brian J. Arnold seeks to find out how the theme of righteousness was received in second century Christianity. A common view is that Augustine was the "fountainhead" of the doctrine of justification, whence it found its way to Luther. The earliest church fathers, on the other hand, are considered to offer very little on the issue. This view is related to the question of whether Paul was used in the second century by the "mainstream" Christians or left in the hands of "heretics", as suggested by the Tübingen school and Adolf von Harnack.[90] Indeed, Paul's strong influence on Valentinians,[91] for example, has been widely

---

[90] Anne Pasquier (2004, 459) has recently stated that "[t]he history of interpretation of Paul's Letters during the second century, it has been said, is essentially the history of Gnostic exegesis."

[91] Unfortunately, Valentinian works in Coptic are beyond the scope of this article.

recognized since the pioneering work by Elaine Pagels.[92] Arnold analyses the idea of justification in Clement of Rome, Ignatius of Antioch, *Epistle to Diognetus*, *Odes of Solomon*, and Justin's *Dialogue with Trypho* and concludes that the alleged gap between Paul and Augustine is one of scholarship and not of sources. According to Arnold, the second-century fathers "believed in justification by faith, and … their view was consistent with the so-called 'traditional' reading of Paul. They tend to present their view of justification forensically and often contrast justification with works-righteousness. Additionally, they use other doctrines associated with Paul, such as imputation and substitutionary atonement."[93] Arnold believes to have proven that "[t]he Apostle was not hijacked by either the proto-orthodox or the Gnostics – he was used by every group despite their varying theological persuasions."[94] Other scholars have also rejected the extreme position of 19th-century scholarship and replaced it with a recognition of Paul's influence on a variety of strands of early Christianity. Importantly, Arnold does not consider the occurrence of διϰ-terminology or explicit references to Paul to be decisive for detecting ideas related to righteousness in the early fathers.[95] Moreover, Arnold makes no claim as to the correct understanding of Paul's righteousness theology or how faithfully the second century fathers reproduce Paul's ideas.[96] Our analysis below suggests significant differences in the treatment of righteousness between these authors.

Nick Needham looks at righteousness language by early church fathers from a wider perspective of the first four centuries. He notes "a very prominent strand of usage in which [justification/righteousness] has a basically forensic meaning" and a more modest strand where the concept is understood as moral transformation and "sanctification".[97] The forensic meaning manifests as "a not-guilty verdict, an acquittal, a declaration of righteousness, a non-imputation of sin, an

---

[92] Especially Pagels 1992.

[93] Arnold 2017, 4.

[94] Arnold 2017, 10.

[95] Arnold 2017, 5, 9. Paul's influence can be detected, for example, in various arguments that use Abraham as legitimation for righteousness, since the argument is originally coined by Paul (Arnold 2017, 11). According to Arnold, "[e]ven though a particular author may not use διϰαιόω or διϰαιοσύνη, the concept of justification may still be present. (...) It is no problem, then, to hunt for justification's tracks even where the word itself is absent." (2017, 5). These contexts might be captured by some word-embeddings methods as well. However, to work properly, these methods require much more extensive textual data than we possess for the second century Christian authors and therefore we cannot employ them for this task here.

[96] Arnold 2017, 1-2.

[97] Needham 2006, 28. Needham takes Clement of Alexandria as a significant exception of the fathers in general, since justification for him almost exclusively denotes "sanctification" or moral conduct (2006, 37).

imputation of righteousness."[98] A good example of forensic thinking is John Chrysostom, who ventures a (markedly anti-Jewish) definition of justification: "What does the word 'justified' mean? That if there could be a trial and an examination of the things He had done for the Jews, and of what had been done on their part towards Him, the victory would be with God, and all the right would be on His side."[99] The forensic and vindicative outlook is evident also in the way the fathers treat "justification" and "condemnation" as equal opposites: "Whom God has condemned, who shall justify?" exclaims Athanasius.[100] According to Needham, the fathers also view justification against other concepts from the forensic domain, such as "guilt". The fathers also identify justification with "forgiveness, remission, pardon, or acquittal." Thus Origen: "If anyone acts unjustly after justification, it is scarcely to be doubted that he has rejected the grace of justification. For a person does not receive the forgiveness of sins in order that he should once again imagine that he has been given a license to sin; for the remission is not given for future crimes, but for past ones."[101]

The forensic outlook, according to Needham, often takes in the early fathers the form of imputation, that is, "reckoning or crediting something to someone's account, a synthesis of legal and financial metaphors, where the books that are being kept are 'judgment books.'"[102] The idea of non-imputation of sins was, however, more common than that of imputation of righteousness.[103] Furthermore, Needham emphasizes that the idea of atonement as penal substitution by Christ is already visible in the fathers, for example, in Justin Martyr, and particularly strongly in John Chrysostom.[104]

As a next step, we look at the early Christian authors in the AGT corpus employing the same distant reading method we employed in the case of comparing with a subselection of pre-Christian authors and subcorpora: comparing cosine similarity of TFIDF vectors based on terms appearing in sentences containing the term δικαιοσύνη. In total, the AGT dataset contains 30 early Christian authors from the first 4 centuries.[105] However, only 26 of them use the term δικαιοσύνη. To

---

[98] Needham 2006, 36.

[99] *Hom. Rom.* 6, on verse 3.4. See Needham 2006, 28.

[100] Athanasius, *Ep. Aeg. Lib.* 19. See Needham 2006, 29.

[101] Origen, *Comm. Rom.* 3.9.4. See Needham 2006, 30.

[102] Needham 2006, 32 offers examples from authors and works such as Clement of Rome, Justin Martyr, Irenaeus, Clement of Alexandria, *Apostolic Constitutions*, Ambrose, John Chrysostom.

[103] Needham 2006, 33.

[104] Needham 2006, 33-35. Unfortunately, John Chrysostom is not included in the AGT corpus.

[105] Unfortunately, the works of John Chrysostom and the *Epistle to Diognetus* are missing in the corpus.

have a broader picture and for the purposes of comparison, we will also include Philo of Alexandria and Josephus Flavius. Thus, we will construct a similarity matrix based on 28 authors or subcorpora.

A matrix of this extent (28 rows and 28 columns) becomes too large to be effectively explored by visual inspection here. Thus, we can inspect this similarity matrix by means of a network instead. For this purpose, the similarity matrix can be treated as an adjacency matrix defining nodes and edges for a network, with individual cell values representing edges' weights (see Figure 5). Since such a network tends to be very dense (zero similarity values are rare) and therefore not very informative, we have to set up a cut off threshold for minimal accepted weight to construct an edge. Figure 5 uses a cut off threshold equal to 0.37.[106]
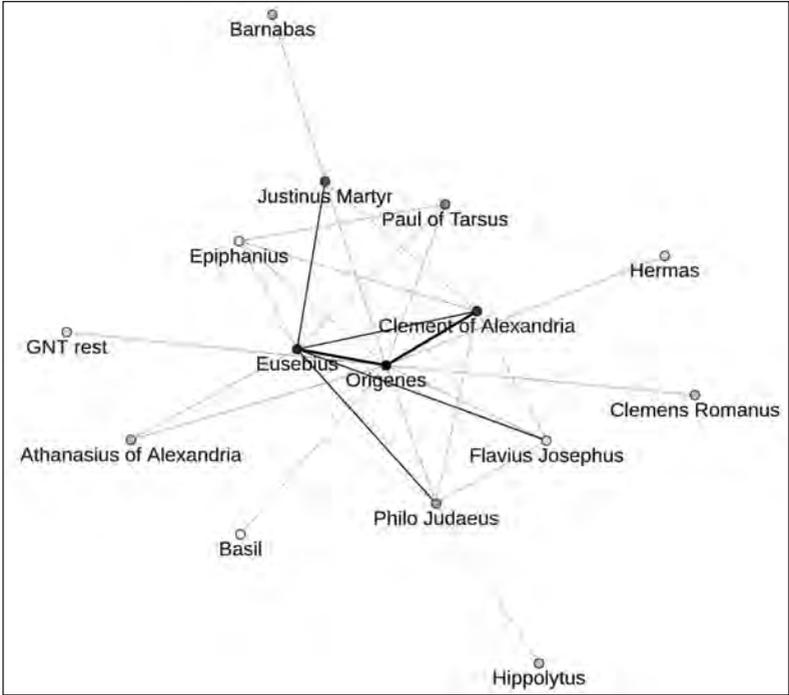

Figure 5. Similarity networks of authors based on cosine similarity of TFIDF vectors.

In the network depicted on Figure 5 we observe Paul is connected to 3 other authors: Epiphanius, Eusebius, and Origen. These are authors with whom he has the highest similarity value based on the TFIDF vectors. In other words, these authors use the term δικαιοσύνη most similarly to Paul. It is striking that all three were active either in the 3rd or 4th century and that Paul is not tied with any first or second century author. This creates distance between Paul and the second century writers, who understood righteousness in a legal sense. Adding the coptic works from Nag Hammadi to the analysis would be relevant for the classic question of "Pauline captivity" by the "heretics". What is further remarkable is the position of Josephus and Philo: both are quite close to the center of the network. In fact, with the cut-off threshold equal to 0.37, Philo maintains the only tie connecting Hippolytus with the rest of the network. Finally, the most central position with the network is occupied by Origen, maintaining ties both to authors preceding him and following him.

IV. Conclusion

The interpretation of righteousness terminology in Paul and surrounding literature is a hotly debated topic of the New Testament and early Christian studies. The contribution of the article to this academic debate is two-fold. First, we repeatedly demonstrated that the distant reading methods return results highly comparable to results achieved by close reading. One example is the close association between righteousness and virtues in the works of Isocrates and Aristotle observed by distant reading but also noted by Charles Lee Irons. The graph of the temporal distribution of δικαιοσύνη is also aligned with Irons's view that this term was frequently used outside Greek biblical texts. These are very relevant results as they validate the applicability of the methodological combination of close and distant reading and reveal the potential of this approach to shed new light on the arguments present in these debates. Secondly, even despite the fact that we did not go very deep into the quantitative results, they already point to interesting trends in the texts that can be further explored in future research into this topic. Especially the future elaboration of the network analysis focusing on the shared righteousness-vocabularies of the words with highest TFIDF among the early Christian authors represents a promising path to study the writing of these authors and their usage of terms associated to righteousness in the contexts of relationships they form on textual level. Despite the significant potential of distant reading, we are aware of its limits with respect to incomplete historical data or difficult lemmatization processes and we do not propose to apply these quantitative tools in methodological separation from close reading, on the contrary, function-

al distant reading requires insights from close reading both in the input phases of the analysis as well as during the interpretation of the results.

For future work, we are also aware of other contexts that can contribute to these textual trends such as the degree of affluence in the society or shifts in meaning related to different cultural translations of righteousness. For example, ancient coins might be a relevant medium for future exploration as they reveal one relevant issue with respect to depicting righteousness. On coins of the Roman Empire, the goddess Dikaiosyne is depicted identically as the Roman goddess Aequitas, that is, as standing and holding scales in one hand and cornucopia in the other.[107] If Roman society perceived and understood Dikaiosyne based on the aspects of Aequitas it would be relevant to explore how this could have affected the use of terms associated with righteousness in the texts analyzed in the present article. The level of affluence in society in the years of an ancient author's writing activities is also a promising context for future academic scrutiny as the academic debate produced hypotheses suggesting that morally loaded terms are more easily upheld and transferred in more prosperous and affluent times.[108] Both these factors (iconography on coins, level of affluence) can be also approached quantitatively in potential future enquiries into this topic. Based on the iconographic interchangeability between Dikaiosyne and Aequitas on Roman coins, we can explore by the methods presented in this article whether there are aspects related to the goddess Aequitas present in the word contexts surrounding righteousness in the works Christian authors active in the era of the Roman empire. With respect to the affluence hypothesis, it is possible to estimate the potential level of affluence in the society based on a variety of economic proxies (e.g., wages, degree of urbanization, population growth, times of infectious diseases, etc.)[109] and then add an attribute to early Christian writers indicating whether the author wrote in years of affluence or crisis. Then it can be observed if this factor could have potentially had an impact on the righteousness vocabulary of these writers and whether it is possible to identify specific tendencies in their vocabulary contrasting with authors writing in times and environments with different levels of affluence.

References

Arnold, Brian J. 2017. *Justification in the Second Century*, Waco, Tx: Baylor University Press.

---

[107] http://numismatics.org/ocre/, https://rpc.ashmus.ox.ac.uk/
[108] Baumard *et al*. 2015.
[109] Ober 2015.

Barr, James. 1961. *The Semantics of Biblical Language*. Oxford: Oxford University Press.

Baumard, Nicolas, and Alexandre Hyafil, Ian Morris, Pascal Boyer. 2015. Increased Affluence Explains the Emergence of Ascetic Wisdoms and Moralizing Religions. *Current Biology*, 25 (1), 10-15.

Blackwell, Christopher W., and Smith, Neel. 2019. The CITE Architecture: a Conceptual and Practical Overview. In: M. Berti, ed. *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution*. Berlin-Boston: De Gruyter, 73-94.

Celano, Giuseppe G. A. 2019. The Dependency Treebanks for Ancient Greek and Latin. In: M. Berti, ed. *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution*. Berlin-Boston: De Gruyter, 279-298.

Cerrato, Lisa *et al*. 2020. PerseusDL/canonical-greekLit 0.0.2711 (Version 0.0.2711). Zenodo (http://doi.org/10.5281/zenodo.4067170).

Crane, Gregory R. 2020. First1kGreek (Version 1.1.5070). Zenodo (http://doi.org/10.5281/zenodo.4091475).

Czachesz, István. 2016. Network Analysis of Biblical Texts. *Journal of Cognitive Historiography* 3, 43-67.

Deerwester, Scott, and Susan T. Dumais, George W. Furnas, Thomas K. Landauer, Richard Harshman. 1990. Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science. American Society for Information Science* 41, 391-407.

Esler, Philip. 1998. *Galatians*. London: Routledge.

Franzini, Greta, and Stefan Jänicke, Muhammad Faisal Cheema, Gerik Scheuermann. 2015. On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In: *Proceedings of the Eurographics Conference on Visualization (EuroVis) (2015)* (https://doi.org/10.2312/eurovisstar.20151113).

Galea, David, and Peter Bruza. 2015. Deriving Word Association Networks from Text Corpora. In: Bruno G. Bara, Gabriella Airenti, Giulio Sandini, eds. *Proceedings of the EuroAsianPacific Joint Conference on Cognitive Science* (EAPCogSci 2015), 252-257.

Gorman, Vanessa B. 2020. Dependency Treebanks of Ancient Greek Prose. *Journal of Open Humanities Data* 6, 1 (https://zenodo.org/record/3596076#.XlZ7CxP7Su4).

Harris, Zellig S. 1954. Distributional Structure. WORD 10, 146-162.

Hill, David. 1967. *Greek Words and Hebrew Meanings: Studies in the Semantics of Soteriological Terms*. Cambridge: Cambridge University Press.

Irons, Charles Lee. 2015. *The Righteousness of God: A Lexical Examination of the Covenant-Faithfulness Interpretation*. Tübingen: Mohr Siebeck.

Jurafsky, Daniel, and James H. Martin. 2017. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (Third Edition Draft)*.

Kaše, Vojtěch. 2021a. LAGT (Version v1.0). Zenodo (http://doi.org/10.5281/zenodo.4971946).

Kaše, Vojtěch. 2021b. tempun (Version v0.1). Zenodo (http://doi.org/10.5281/zenodo.4650989).

Kaše, Vojtěch, and Nina Nikki, Tomáš Glomb. 2021. ECCE_DIK (Version v1.2). Zenodo (http://doi.org/10.5281/zenodo.4972801).

Kluyver, Thomas, and Benjamin Ragan-Kelley, Fernando Pérez, Brian Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica Hamrick, Jason Grout, Sylvain Corlay, Paul Ivanov, Damián Avila, Safia Abdalla, Carol Willing, Jupyter Development Team. 2016. Jupyter Notebooks – a Publishing Format for Reproducible Computational Workflows. In: Fernando Loizides, Birgit Schmidt, eds. *Positioning and power in academic publishing: players, agents and agendas: proceedings of the 20th International Conference on Electronic Publishing*. Amsterdam: IOS Press, 87-90.

Landauer, Thomas K., and Danielle S. McNamara, Simon Dennis, Walter Kintsch, eds. 2007. *Handbook of Latent Semantic Analysis*. New York-London: Routledge.

Lenci, Alessandro. 2018. Distributional Models of Word Meaning. *Annual Review of Linguistics* 4, 151-171.

Mihalcea, Rada, and Dragomir Radev. 2011. *Graph-Based Natural Language Processing and Information Retrieval*. Cambridge: Cambridge University Press.

Mikolov, Tomas, and Kai Chen, Greg Corrado, Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. *arXiv [cs.CL]*. Retrieved (http://arxiv.org/abs/1301.3781).

Moretti, Franco. 2013. *Distant Reading*. London-New York: Verso.

Needham, Nick. 2006. Justification in the Early Church Fathers. In: Bruce L. McCormack, ed. *Justification in Perspective: Historical Developments and Contemporary Challenges*. Grand Rapids: Baker Academic, 25-53.

Nikki, Nina. 2019. *Opponents and Identity in Philippians*. Leiden: Brill.

Ober, Josiah. 2015. *The Rise and Fall of Classical Greece*. Princeton: Princeton University Press.

Pagels, Elaine. 1992. *The Gnostic Paul: Gnostic Exegesis of the Pauline Letters*. New York: Trinity Press International.

Painter, John. 2008. *1, 2, and 3 John*. Collegeville, Mn.: Liturgical Press.

Palola, Jukka. 2011. *Vanhurskautta, oikeutta vai uskollisuutta?* דק *ṣdq-sanue vuoden 1992 Kirkkoraamatun Psalmien kirjassa*. Dissertation, University of Helsinki.

Pasquier, Anne. 2006. The Valentinian Exegesis. In: Charles Kannengiesser, ed. *Handbook of Patristic Exegesis: The Bible in Ancient Christianity*. Leiden: Brill, 454-470.

Sanders, Ed P. 1977. *Paul and Palestinian Judaism: A Comparison of Patterns of Religion*. Philadelphia: Fortress Press.

Schmid, Hans Heinrich. 1968. *Gerechtigkeit als Weltordnung: Hintergrund und Geschichte des alttestamentlichen Gerechtigkeitsbegriffes*. Tübingen: Mohr.

Schweitzer, Albert. 1998. *The Mysticism of Paul the Apostle*. Transl. by William Montgomery. Baltimore: Johns Hopkins University Press. (1st ed. 1931, 2nd ed. 1953).

Seifrid, Mark A. 2001. Righteousness Language in the Hebrew Scriptures and Early Judaism. In: D.A. Carson, Peter T. O'Brien, Mark A. Seifrid, eds. *Justification and Variegated Nomism. Vol. 1. The Complexities of Second Temple Judaism*. Tübingen, 415-442.

Seifrid, Mark A. 2004. Paul's Use of Righteousness Language against Its Hellenistic Background. In: D.A. Carson, Peter T.O'Brien, Mark A. Seifrid, eds. *Justification and Variegated Nomism. Vol. 2. The Paradoxes of Paul*. Tübingen, 39-74.

Špiclová, Zdeňka, and Vojtěch Kaše. 2020. Distant Reading of the Gospel of Thomas and the Gospel of John: Reflection of Methodological Aspects of the Use of Digital Technologies in the Research of Biblical Texts. *Open Theology* 6, (1), 423-439. (https://doi.org/10.1515/opth-2020-0111).

Strang, Gilbert. 2016. *Introduction to Linear Algebra* (5th ed.). Wellesley: Wellesley-Cambridge Press.

Underwood, Ted. 2017. A Genealogy of Distant Reading. *Digital Humanities Quarterly* 11, 1-12.

Van Rossum, Guido, and Fred L. Drake. 2009. *Python 3 Reference Manual*. Scotts Valley: CreateSpace.

Wrede, William. 1907. *Paul*. Transl. by E. W. Lummis. London: Philip Green (1st ed. 1904)

Ziesler, John A. 1972. *The Meaning of Righteousness in Paul: A Linguistic and Theological Enquiry*. Cambridge: Cambridge University Press.

Vojtěch Kaše
Department of Philosophy
Faculty of Philosophy and Arts
at the University of West Bohemia in Pilsen
Sedláčkova 19, 306 14 Plzeň, Czech Republic
kase@kfi.zcu.cz
School of Culture and Society, History
Aarhus University
Jens Chr. Skous Vej 5, 8000 Aarhus, Denmark
vojtech.kase@gmail.com

Nina Nikki
Department of Biblical Studies
Faculty of Theology
Vuorikatu 3, 00014 University of Helsinki, Finland
nina.nikki@helsinki.fi

Tomáš Glomb
Department for the Study of Religions
Faculty of Arts, Masaryk University in Brno
Arna Nováka 1/1, 602 00 Brno, Czech Republic
tomas.glomb@gmail.com