The Un-Platonic *Menexenus*: A Stylometric Analysis with More Data

The *Corpus Platonicum* is one of the most well-known and most influential works of ancient literature. Yet, it still has unresolved challenges regarding authorship of some of the works. Since the middle of the nineteenth century, stylometry, that is the counting and statistical analysis of features in the language of an author, has been applied to the *Corpus Platonicum* in order to date works by Plato, prove or disprove the authenticity of his work, or to analyze how he used different styles; for instance, to characterize different people in his dialogues.[1] This article revisits the second of these applications with a special emphasis on the question of the authorship of the *Menexenus*.

The results of previous stylometric analyses of the *Corpus Platonicum* were, with regards to the question of authorship, inconclusive.[2] One reason for this probably lies in the selection of features—that is, the frequency of certain words, the number of words in a sentence, the quantities of *clausula*, or any other element of the text that can be observed and counted and that is suspected to be part of the stylistic fingerprint of the author—and the number of those features that had been extracted in prior analyses.[3] In the first non-computational stylometric analyses the number of extracted features was very low: Campbell and Dittenberger each counted eight lexical features and Billig used fourteen metric features.[4] One of the most extensive non-computational stylometric analyses was performed by Ritter in 1910; he used forty-three linguistic features.[5] Unfortunately, their studies only help a little to evaluate the authenticity of the *Menexenus*: Campbell did not focus on the *Menexenus*, because the aim of his study was the analysis of the *Sophistes* and the *Politicus*; Billig did not include all of the *Corpus Platonicum* and one of the dialogues that was missing was the *Menexenus*; Dittenberger did not include it because it was considered unauthentic by most scholars in Germany at the time; and while Ritter saw similarities between the *Menexenus* and the *Symposium*, his research

---

[1] A good example of the latter has been published in issue 58 of this journal. See D. Sansone, "Stylistic Characterization in Plato: Nicias, Alcibiades, and Laches," *GRBS* 58.2 (2018) 156–176.

[2] A more complete survey than can be given in this focused article was undertaken by L. Brandwood, "Stylometry and Chronology," in R. Kraut (ed.), *Cambridge Companion to Plato* (Cambridge 1992) 90–120.

[3] Or as J. Burrows, "'Delta'—A measure of stylistic difference and a guide to likely authorship," *Literary and Linguistic Computing* 17.3 (2002) 267–287, 268, put it, "in this sort of work on language, so our researches teach us, a wealth of variables, many of which may be weak discriminators, almost always offer more tenable results than a small number of strong ones."

[4] See L. Campbell, *The Sophistes and Politicus of Plato* (Oxford 1867) I–XLV; W. Dittenberger, "Sprachliche Kriterien zur Chronologie der platonischen Dialoge," *Hermes* 16 (1881) 321–345; and L. Billig, "Clausulae and Platonic Chronology," *JP* 35 (1920) 225–256.

[5] C. Ritter, *Untersuchungen über Plato. Die Echtheit und Chronologie der platonischen Schriften* (Stuttgart 1888). Also see C. Ritter, *Platon: Sein Leben, Seine Schriften, Seine Lehre* (München 1910), esp. 232–272.

only included works in the *Corpus Platonicum* and therefore lacked comparison texts.[6] In addition, Ritter admitted that the number of features he observed in the *Menexenus* was low.[7]

In one of the first computational stylometric analyses of the *Corpus Platonicum*, Ledger, in his book *Re-Counting Plato* (1989), used thirty-seven lexical features as well as comparison texts. In his study, Ledger found unusual results for the *Menexenus* compared with other Platonic works; however, he didn't trust those results since he saw the authenticity of the *Menexenus* as confirmed by references by Aristotle that may point to the *Menexenus*.[8] In fact, he excluded the *Menexenus* from his authenticity test that would have treated the work as its own author, writing that this was "partly through oversight, partly because what I am attempting here is illustrative rather than definitive, and partly because I believe them to be genuine."[9] In addition, the computational analysis of Greek works was more difficult thirty years ago. Given the challenges of encoding, and therefore also counting Greek words in the 1980s, Ledger defined features based on the first and last letters of a word, arguing that this would be the best way to observe lexical features of an inflected, morphologically complex language when limited by processing and memory power.

Empowered by technological progress, however, the last decade of Natural Language Processing research has developed promising automated analytical methods to process and classify vast amounts of texts. For instance, the R-Package *stylo* makes it easy to extract lexical features, such as the frequency of words or the frequency of character n-grams; that is, the frequency of a sequence of n characters (e.g. a 4-gram consists of four characters).[10] Both feature sets—frequency of words and character n-grams—are complimentary to each other. While the frequency of words almost entirely focuses on function words (that is, words with little semantic meaning that still mark a stylistic decision by the author, such as καὶ, δὲ, μὲν, τε), measuring the frequency of sequences of four characters will also show, for instance, an author's preference for certain compounds (e.g., περὶ, περίπλεω), certain endings, or even sounds.

Furthermore, and equally empowered by technological progress, a lot of machine-actionable textual data has been produced by the Perseus Digital Library and the Open Greek

---

[6] Dittenberger, "Sprachliche Kriterien" 322 n.1.; Ritter, *Platon* 496.

[7] Ritter, *Untersuchungen* 97–98. Also cf. 125, where he assumes that the noticeable change in content in the *Menexenus* might be evidence that Plato became happier again: "Auch kann ich mich nicht entschliessen den Menexenus als unecht zu verwerfen, der wegen seiner leichten Haltung in die zweite Hälfte meiner ersten Periode nicht hereinpassen will. Ich muss so annehmen, dass … eine heitere und freiere Gemütsstimmung sich bei Plato herstellte."

[8] G. R. Ledger, *Recounting Plato: A Computational Analysis of Plato's Style* (Oxford 1989) 163.

[9] Ledger, *Recounting Plato* 105.

[10] M. Eder, J. Rybicki, and M. Kestemont, "Stylometry with R: A Package for Computational Text Analysis," *R Journal* 8.1 (2016) 107–121.

and Latin Project (OGLP), whose staff has digitized and curated not only the preponderance of the *Corpus Platonicum* and most of the Greek authors commonly read in schools and universities, but also, in collaboration with CHS,[11] in what is called the *First Thousand Years of Greek Project*, digitized extant works of the first one thousand years of Greek literature.[12] Now, for the first time in history, it is possible to use computational analysis for the preponderance of Ancient Greek literature.

Using more data also enables a better understanding of Platonic style. While it is difficult to compare Platonic works just with themselves, it is possible to detect features with greater discriminatory power when we compare as many works as possible. This observation is not new, however: in 1989, Ledger already used not only works included in the *Corpus Platonicum*, but also other prose works of the fifth and fourth century BCE by Aeschines, Isaeus, Isocrates, Lysias, Thucydides, and Xenophon. Yet, Ledger admitted that his experiment ignored the component of corruption in the process of textual transmission. While I agree that given the amount of data some individual corruptions would not greatly affect the overall result, by preselecting only texts from the fourth and fifth centuries, Ledger additionally ignores— without mentioning it—the possibility of a later forgery of individual works. Since we now have more data, I ran stylometric experiments without preselecting a certain style of Ancient Greek. The machine can now easily cluster—that is, put the works into groups that have statistically similar features—bigger groups by itself and we do not run the risk of influencing the results through selection bias a priori.

Another advantage of more data is that morphological normalization—that is, reducing each word to its dictionary entry—is less urgent. The problem with analyzing morphologically complex languages with statistical tools is essentially a feature-frequency problem; that is, a feature will be, virtually speaking, ignored if it occurs only a statistically irrelevant number of times. In Ancient Greek, essentially the same lexical unit can have more than ten different morphological appearances. For the computer, all of those different appearances are entirely different features: 'Πλάτωνα' has the same difference from 'Πλάτων' as it has from 'Εὔβουλος' if we choose to select the frequencies of morphologically unnormalized words for our statistical stylometric analysis. But by including more data, the most frequent words are less prone to becoming statistically irrelevant. In fact, even the absence of a word in a work becomes a statistically relevant feature if we can otherwise find it among the most frequent words in the

---

[11] Author identifying footnote removed for peer review.
[12] For this part of the corpus see http://opengreekandlatin.github.io/First1KGreek/. The Perseus text can be retrieved at https://github.com/PerseusDL/canonical-greekLit. Most of the texts are also accessible using the Scaife viewer: https://scaife.perseus.org.

corpus. Additionally, measurements of the frequency of 4-character sequences are also less affected by this kind of feature-frequency problem. Both kinds of features were independently used in my experiments.

In what follows, I will describe a thorough series of stylometric experiments on the machine-actionable corpus of Ancient Greek text produced by Perseus, OGLP, and the CHS and will focus on one finding, of potentially many more findings to come. That finding is that although the *Menexenus* was transmitted in the Platonic canon of the tetralogies, there is very little stylistic similarity with the rest of the *Corpus Platonicum*. While this, in itself, does not necessarily prove that the *Menexenus* wasn't written by the authorial entity identified as Plato, it certainly indicates this if its inclusion cannot be strongly supported by philological or philosophical arguments.

Data and Method

This experiment used the corpus generated by Perseus, OGLP, and CHS as of December 2017. This corpus includes CTS identifiers for all works and workgroups. For instance, tlg0059 stands for the *Corpus Platonicum* and tlg0059.tlg028 for the *Menexenus*. At the time of research, the corpus contained almost thirty million words in 1,044 works. Because not all of those are entirely in Ancient Greek (e.g., Perseus includes some translations in their repository), I have filtered out all non-Greek works, which reduced the corpus to twenty-three million words, and I then reduced this further by only including works with more than 1,500 words of Ancient Greek.[13] This resulted in 825 works with a total of 23,113,247 words. However, this also included some works with more than one version, because Perseus digitized more than one version or the modern language translation still contained more than 1500 Greek words. For those instances I chose the version that contained more words. In the end, the data used for the stylistic experiments comprised 800 works that together contained around 22,250,000 words, which makes this research the biggest stylometric experiment for Ancient Greek at the time of writing.[14]

From this corpus I extracted twenty-five different feature sets. A feature set is the frequency of certain features in the corpus. Authorship can be deduced from the preferences of

---

[13] At a talk at DH2017, M. Eder suggested that a minimum of 2000 words be used for English. I set the threshold to 1500 so as to not exclude Plato's *Clitopho*. I have since repeated the experiments by setting higher thresholds of 2000 and even 3000 words but the results do not change regarding the *Corpus Platonicum*. See M. Eder, "Short Samples in Authorship Attribution: A New Approach," paper given at the DH2017 conference in Montreal. Abstract accessible at: https://dh2017.adho.org/abstracts/341/341.pdf.

[14] I have made the data and the R-programming script that wrangled and analyzed the data available under open licenses at GitHub (MIT for the script and CC-BY for tables and visualizations), so readers can reproduce the experiment or expand on it. For the blind peer-review:
https://drive.google.com/drive/folders/1wvmMNndl6utRCVqNvcPbtSxqdp-i_om-?usp=sharing.

authors for certain words (or other features). The feature sets in my experiments consisted either of words or 4-grams in the corpus. The twenty-five feature sets could themselves be divided into five different groups. First, was a group solely based on word frequency, in which the 1000, 500, 300, 200, and 100 most-frequent words were used as features. Second, was a group of the 1000, 500, 300, 200, and 100 most-frequent 4-grams. Both of those feature sets were extracted using the R library *stylo*.[15] In the third group the words were ranked according to their frequency in the corpus, but were excluded if they did not occur in the *Menexenus*. This essentially gave every feature that does not occur in the *Menexenus* the weight 0; that is, we assume for the experiment that a word not occurring in the *Menexenus* is not important or its importance cannot be assessed and it should be excluded. It remains an open question though, whether that assumption is correct or not, but I nevertheless allowed the bias of the experiment to slightly favor detecting a weak Platonic signal in the *Menexenus*. This group and the fourth and fifth groups have been used to double-check the clear result (that is, that the *Menexenus* does not seem Platonic in style) of the experiment run with the first two groups. The fourth and fifth groups consist of two sets, one of words and one of 4-grams, each of which was built from the sets generated by *stylo* in groups one and two.

To produce the fourth and fifth group, machine learning was used to prune the feature sets. Put simply, because the CTS workgroups often represent works by one author (e.g., tlg0007 is Plutarch), we can use this bibliographic information to train algorithms with a programme such as the Data Mining software WEKA.[16] This means that we divide the corpus into training and test set(s),[17] then train the algorithm in such a way that it is optimized for the task of classifying a work as belonging to a workgroup on the basis of the frequencies of the features. For instance, if we have the work tlg0007.tlg001, which is Plutarch's *Life of Theseus*, we can train the algorithm so that when it looks at word frequencies in that work it can determine that this work belongs to the workgroup tlg0007 (works by Plutarch). In WEKA, we can train algorithms to classify the data or use out-of-the-box classifiers to see how well they perform on the desired classification task. For instance, one algorithm in WEKA—called ZeroR—always guesses the most frequent class in training set(s); that is, it no matter which work you feed into the algorithm it always guesses Libanius (tlg2200), the author with the most works in our corpus. Of course, this is not a very smart algorithm, but it provides a very nice

---

[15] Eder et al., "Stylometry" 107–121.

[16] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, "The WEKA Workbench," online appendix for M. Kaufmann (ed.), *Data Mining: Practical Machine Learning Tools and Techniques* (2016). Available at: https://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf.

[17] Since I used 10-fold cross-validation when training the algorithms, the plural 'sets' is more correct, but the principle relies on one test set and one training set.

baseline to test how effective our algorithms are: ZeroR achieved an accuracy of 12.13%. OneR is another simple algorithm in WEKA. It attempts to find the most discriminative feature—that is, the frequency of one word or 4-gram on the basis of which alone most correct classifications in the training set could be made—in the dataset and uses the value of that feature as the rule to determine the class. When I used OneR on 4-grams the feature was 'α ì _ ἐ,' which achieved 22.88% accuracy, and when I used it on words the feature was 'μὲν,' which achieved 20%. While these do not have a great accuracy either, it already means that by simply measuring the mean frequency of the word 'μὲν' and building a rule from this, we can correctly identify the author of an Ancient Greek work in one out of five cases.

Two slightly more complex algorithms perform a lot better on the data: J48 and Naïve Bayes. J48 is a classifier that builds a decision tree based on the training data and then uses that decision tree to classify the test data.[18] Using the stylometric feature sets, the J48 classifier correctly classified 50.75% of the instances when using character 4-gram frequencies and 50.5% when using word frequencies. Naïve Bayes is a probabilistic classifier that, for each feature attribute, computes as many probabilities as there are classes and then chains them to make a classification.[19] Using the stylometric feature sets, the Naïve Bayes classifier correctly classified 65.88% of the instances when using character 4-gram frequencies and 64.75% when using word frequencies.

While WEKA allows us to determine how well each trained algorithm performs and thus generate a baseline that our stylometric specific measures must beat, we can also use what we learned from analyzing the classifiers for feature selection. In WEKA this is called 'Attribute Selection.' For instance, if we look at the results of the J48 and Naïve Bayes classifier for the *Corpus Platonicum* (tlg0059), we see that while the overall accuracy of Naïve Bayes when using word frequencies was higher than the one for J48, the numbers for Plato were worse (Naïve Bayes only had a precision of 0.4 for the class tlg0059, while J48 had a precision of 0.8). With the *Ranker* method in WEKA, we can rank the attributes according to their usefulness. In the two sets built with Weka, I ranked the feature sets that were extracted by *stylo* and removed the attributes that *Ranker* evaluated negatively. For the word frequency *Ranker* evaluated the attributes using J48, while for the 4-gram frequencies it used Naïve Bayes because of their higher precision when classifying Platonic works in each feature set.

I should stress that I had initially only used the most frequent words and most frequent 4-grams; however, I later decided to include more feature sets because the *Menexenus* did not

---

[18] Witten et al., "WEKA Workbench," 92.
[19] Witten et al., "WEKA Workbench," 34.

seem to cluster with Plato when using the original sets. My aim was to calibrate the features to pick up on a potentially weak *Menexenus* signal to see if there is any evidence at all for the hypothesis of an authorship of the *Menexenus* by Plato. The feature sets three, four, and five were my attempt to defend this hypothesis. Not surprisingly, the WekaWord feature/measure detected similarities between the *Menexenus* and the *Corpus Platonicum* best.

Measuring the frequency of each feature resulted in a large table, where the rows are the unique identifiers of the works and the columns are the names of the measured features. Given that in our experiments we observe up to 1000 features (either 1000 words or 1000 n-grams) at once and we looked at 800 ancient works, the biggest matrix has 800,000 cells. It is clear that this would be too much data for manual qualitative analysis. Nor can we use graphs to express the data, because they are essentially discrete sets: although some connections between frequencies of individual words exist (e.g., μέν—δέ), the frequency of 'ἀνθρώπους,' for instance, has no direct influence on the frequency of 'καί.' Each data point, or work, is best thought of as a set of one thousand values. Given that the order of those values is always the same for each data point, we can think of each work as a point in a high-dimensional coordinate system, where each dimension is the frequency of a word or n-gram. Thus, when using the frequencies of the most-frequent 1000 words as features in our stylometric analysis, we create a 1000-dimensional space in which each work is a single point in that space, represented by its set of 1000 values. The challenge is that high-dimensional spaces are difficult for the human brain to visualize, imagine, and interpret.

In order to make sense of the data, we must address this challenge. In the experiments I pursued two independent approaches: one is a t-SNE dimensionality reduction approach—that is, simplifying the 1000-dimensional space to a 2-dimensional space—that then enables us to use simple Euclidean distances to cluster the data, and the other measures several types of distances between points directly in the high-dimensional space and then use those distances to cluster the data. The same clustering method—that is, grouping the works according to their data similarity—has been applied to the data. The clustering used Ward's method; to be precise, the ward.D-implementation of the programming language R.[20] Ward's method is an agglomerative clustering method; that is, it produces a tree by combining close clusters with each other until there is nothing to join. Because we know in which step each cluster was created we can easily cut any number of clusters out of that tree, where the minimum number of clusters we can cut is one and the maximum is the number of data points (in our data-set that is 800).

---

[20] F. Murtagh and P. Legendre, "Ward's Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm," arXiv:1111.6285v2 [stat.ML] (11 Dec 2011), https://arxiv.org/pdf/1111.6285v2.pdf.

Because our data was grouped by humans into 153 workgroups representing something like an author, I cut the tree for automatic evaluation into 153 clusters. Additionally, I also looked at the whole trees.[21]

The abbreviation t-SNE stands for t-Distributed Stochastic Neighbor Embedding and is a popular algorithm for dimensionality reduction.[22] Essentially, t-SNE attempts to visualize data points that have close proximity in the high-dimensional space as having a close proximity in a low-dimensional space. For example, if we have a high-dimensional data set that we want to visualize in two dimensions, we can observe the t-SNE characteristic groups of data points; that is, the data that is similar to each other looks like they build an 'island' on a map. If we want to visualize the same data set in three dimensions, we would observe 3-dimensional bodies of data points vaguely resembling 'blobs.' Those islands or blobs are essentially visual clusters of data that is similar to each other. If a data point is outside the cluster, it is so different in the high-dimensional space from the rest of the data that it is visually not included on the same island.

Aside from combining t-SNE with two-dimensional Euclidean distances, I have also used distances directly with the high-dimensional data. Because we do not know what kind of space the feature set creates, I have used eight different distances to cluster the data: the Euclidean distance, which is the direct connection between points in a Euclidean space (an example is that which most humans understand as physical space); the Jaccard Distance, which measures the similarities between sets; Cosine Distance and Cosine Delta, which are two distances based on the cosine of the two vectors expressed through the values of each data point; and finally, four other distances that were developed for authorship attribution: Burrows' Delta, Eder's Simple, Eder's Delta, and Argamon's Delta.

Although Eder's Delta, Burrows' Delta, and especially Cosine Delta have been considered very reliable for authorship attribution, the literature disagrees about whether these measures are the best for author attribution in all languages, because there may be language-dependent factors that influence the result.[23] For instance, the best measure for German might

---

[21] These can be accessed at Github, and for peer review:
https://drive.google.com/drive/folders/1wvmMNndl6utRCVqNvcPbtSxqdp-i_om-?usp=sharing.

[22] L. v.d. Maaten, G. Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Research* 9 (2008) 2579–2605. Available at: http://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf.

[23] See J. Rybicki and M. Eder, "Deeper Delta Across Genres and Languages: Do We Really Need the Most Frequent Words?," *Literary and Linguistic Computing* 26.3 (2011) 315–321. For a ranking of the algorithms based on English, French, and German novels see F. Jannidis, S. Pielström, C. Schöch, and T. Vitt, "Improving Burrows' Delta – An Empirical Evaluation of Text Distance Measures," paper given at the DH2015 conference in Sydney. Abstract accessible at:
http://dh2015.org/abstracts/xml/JANNIDIS_Fotis_Improving_Burrows__Delta___An_empi/JANNIDIS_Fotis_Improving_Burrows__Delta___An_empirical_.html. For a better understanding of Delta authorship attribution

not be the best measure for Greek when the most frequent words are used, which is why I used and evaluated eight different measures. In order to evaluate the measures, each feature-set/measure combination has been tested and scored.

Testing the Measures

In order to test the 225 feature-set/measure combinations, I once again used the fact that the CTS URN of each work indicates to which author/workgroup the individual work is attributed. Digital archivists have ordered the 800 works into 153 workgroups. This was not done in a single metadata sprint; rather, it happened over decades based on decisions influenced by traditional assumptions and ease of collection use. As a result, a workgroup sometimes contains more than one author (e.g., the workgroup tlg0013 stands for *Homeric Hymns*), and sometimes there will be—given the complexity of textual transmission—authorship attribution errors. But since a group of librarians, archivists, and philologists involved in the process have decided on 153 workgroups, we can test what a machine could do if its task were to cluster the corpus into 153 groups using a selection of different measures and then compare the results with the human categorization for an estimate of how well the individual measures performed.

If works are part of a cluster that contains works by only one author, we can then consider them correctly classified, whereas if works are in a mixed group, we consider them as classified to the author that holds the majority in that group, while if there is no majority author or if they are a minority author in the group then we consider them misclassified. We then measure the percentage of works that pass the test as classified correctly and thereby get an estimate of how well the feature-set/measure combination performs in an authorship attribution task in a real-world corpus. Since the score is based on membership to a cluster and because it does not really measure the precision in a controlled environment where we are 100% certain of the authorship of a work, but only a value of how well it performed compared to clusters built by humans, I call this the k-score. Although it is theoretically possible to reach a k-score of one, it is very unlikely for two reasons: first, we can assume that not all works are correctly attributed to an author when preparing the URNs, and second, each feature-set/measure combination may be distracted by different signal noise. In FIGURE 1 you can see the distribution of the k-scores of the individual measures. They look very promising; especially if we compare them to the classification scores in Weka mentioned above.

One non-humanist, but stylometric research result of this test is that, although aside from Argamon's Delta all delta measures performed better than all non-authorship attribution

---

measures see S. Evert and T. Proisl, "Understanding and Explaining Delta Measures for Authorship Attribution," *Digital Scholarship in the Humanities* 32.2 (1 December 2017) ii4–ii16. https://doi.org/10.1093/llc/fqx023.

specific measures, Eder's Simple not Eder's Delta performed most precisely and most reliably of them all. This is counter-intuitive, given that Cosine Delta outperforms Eder's Simple for English, German, and French, and can only be explained by the fact that the Latinist and stylometry expert Eder built his measures for authorship attribution for morphologically more complex languages.[24] That being said, it is worth noting that the simple measure outperforms the Delta measure in Ancient Greek, which is a lot more morphologically complex than Latin, and that all measures performed generally well.[25]
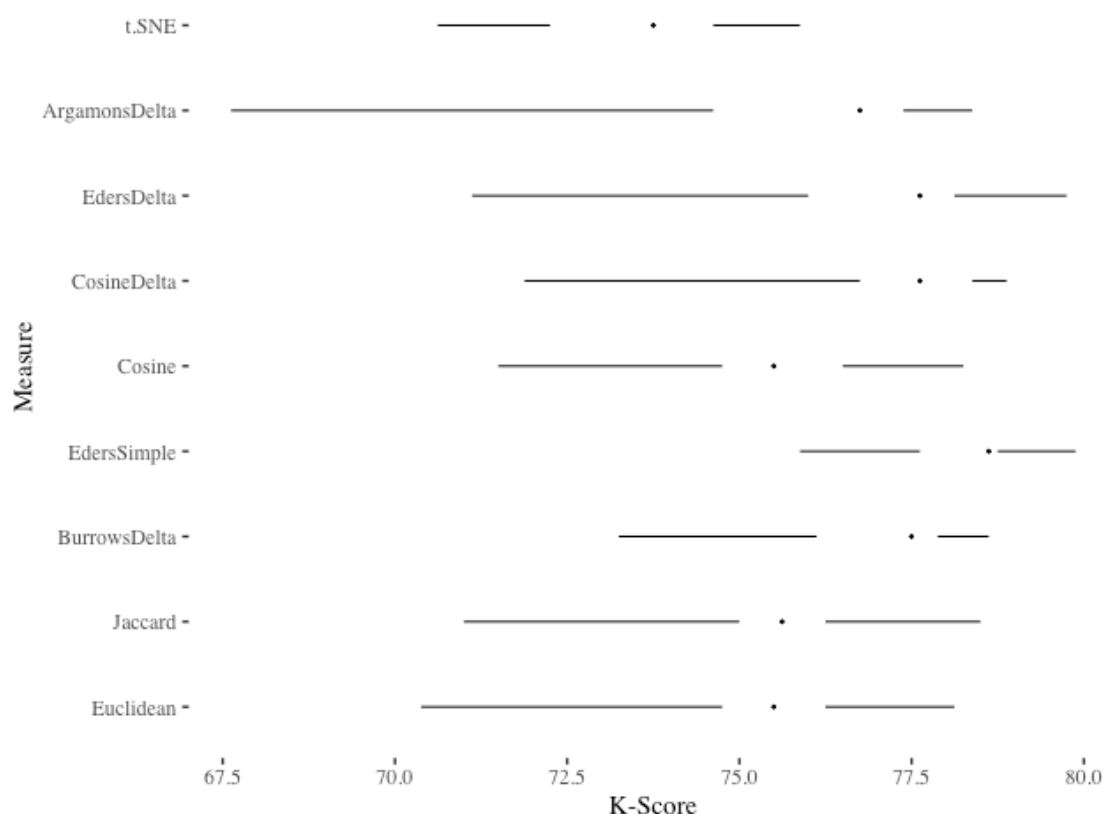

FIGURE 1: Tufte-style box-and-whiskers plot of the k-scores of the measures used in the experiment. The dot shows the median value of all k-scores for the twenty-five different feature sets and the whiskers show the range of the value. Eder's Simple is the most consistent, has the highest individual k-score, and the highest median.

Given that the all feature-set/measure combinations perform mostly between 0.75 and 0.8, we can consider any of the measures as reliable, although Eder's Simple can be considered the most reliable. However, given that each feature-set/measure combination might be affected

---

[24] Cf. F. Jannidis et al., "Burrows' Delta," and S. Evert, T. Proisl, F. Jannidis, S. Pielström, C. Schöch, and T. Vitt, "Towards a Better Understanding of Burrows's Delta in literary authorship attribution," *Proceedings of NAACL-HLT Fourth Workshop on Computational Linguistics for Literature* (Denver 2015) 79–88.

[25] To confirm Eder's Simple as the best measure currently available for Ancient Greek works, this should be followed up by a stylometric experiment based solely on works where we are 100% certain of authorship to get an understanding which method performs better in a more controlled sample than in a real-world data-set.

by different kinds of noise in the data-set (for instance, a strong genre signal) and given that all measures score comparably well, I have compared the results of all 225 feature-set/measure combinations to analyze whether the *Menexenus* might have been misclassified as a work written by Plato.

Results of the Statistical Stylometric Analysis

When using the methods based on t-SNE, the *Menexenus* is well outside the cluster of Platonic works in all visualizations.[26] It doesn't cluster with other works; rather, it sits somewhere close to the works of Lucian and Plutarch and sometimes even Galen. Aside from the *Menexenus*, two other works often appear outside the main cluster of Platonic works: the *Timaeus* and the *Critias*. Those two works almost always cluster with each other and are at times in very close proximity to the main cluster. At times, the *Clitopho* is also outside the main cluster.

Similar observations can be made when we look at the results of the other method, the cluster analysis of all distances: the *Menexenus* rarely clusters with Plato (in only around 10% of the feature-set/measure combinations). In a further five cases it clusters alone, but after further examination of the hierarchical tree, it can be counted to a cluster of other works in the *Corpus Platonicum*, because the parent cluster mainly consists of works from that corpus. The *Menexenus* also rarely clusters with more than two Platonic works at once, often doing so with just one other Platonic work, as shown in TABLE 1.

TABLE 1: Showing how many other works of the *Corpus Platonicum* cluster with the *Menexenus* in each feature-set/measure combination. Five cases where the *Menexenus* clusters alone, but is still part of bigger Platonic clusters are not specifically marked in this table, but

---

[26] Given their size and resolution it is not practical to include those visualizations in a print publication. However, they can be accessed at Github and for peer review here: https://drive.google.com/drive/folders/1wvmMNndl6utRCVqNvcPbtSxqdp-i_om-?usp=sharing .

were included in the analysis. Those are: Burrows' Delta at 300MenexMFW, Eder's Simple at 1000 and 500WekaWord, Eder's Delta and Argamon's Delta at 500 MenexMFW.

	Euclidean	Jaccard	BurrowsDelta	EdersSimple	Cosine	CosineDelta	EdersDelta	ArgamonsDelta	t.SNE
1000MFW	0	0	0	0	0	0	0	0	0
500MFW	0	0	4	0	0	0	0	0	0
300MFW	0	0	0	0	1	5	0	0	0
200MFW	0	0	0	0	0	1	0	0	0
100MFW	0	0	0	1	0	1	0	0	0
1000MF4Gram	0	0	0	0	0	0	0	0	0
500MF4Gram	0	0	0	0	0	0	0	0	0
300MF4Gram	0	0	0	0	0	0	0	0	0
200MF4Gram	0	0	0	0	0	0	0	0	0
100MF4Gram	0	0	0	0	0	0	0	0	0
1000WekaWord	0	0	0	0	0	4	0	0	0
500WekaWord	0	0	0	0	0	0	0	7	0
300WekaWord	0	0	0	1	1	0	0	0	0
200WekaWord	1	0	0	2	0	2	0	1	0
100WekaWord	0	1	2	1	1	4	0	5	0
859Weka4Gram	0	0	0	0	0	0	0	0	0
500Weka4Gram	0	0	0	3	0	0	0	0	0
300Weka4Gram	0	0	0	0	0	0	0	0	0
200Weka4Gram	0	0	0	0	0	0	0	0	0
100Weka4Gram	0	0	0	0	0	0	0	0	0
810MenexMFW	0	0	0	0	0	0	0	0	0
500MenexMFW	0	0	0	0	0	3	0	0	0
300MenexMFW	0	0	0	1	0	0	0	1	0
200MenexMFW	0	0	0	0	0	5	0	0	0
100MenexMFW	0	0	0	0	0	1	0	0	0

While this does not speak strongly for an attribution of the *Menexenus* to Plato, the result becomes even more decisive if we compare the scores shown in TABLE 1 with the individual scores of the feature-set/measure combinations in TABLE 2. The cases where the *Menexenus* clusters with Plato often have lower k-scores and there is no case of the *Menexenus* clustering with another Platonic work when any of the five most frequent 4-gram sets were used (the feature sets with the highest mean k-score). Not one of the top ten highest-scored feature-set/measure combinations has a cluster that includes at least one other Platonic work and the *Menexenus*.

TABLE 2: K-Scores by Feature-Set/Measure Combination. The scores where the *Menexenus* clusters with another work of the *Corpus Platonicum* have a grey background.

	Euclidean	Jaccard	BurrowsDelta	EdersSimple	Cosine	CosineDelta	EdersDelta	ArgamonsDelta	t.SNE
1000MFW	76.88	76.25	77.75	78.88	76	78.12	78.25	76.88	74.38
500MFW	76.5	75.88	77.12	78.75	75.5	78.62	77.25	77.5	74
300MFW	76.12	75.5	77	78.12	75.75	78.5	78.62	76.5	73.38
200MFW	75.88	75	77.88	78.75	75.62	77.62	78.62	77.38	73.38
100MFW	75.38	74.62	78.62	79	75.25	77.75	77.25	76.88	72.12
1000MF4Gram	77.75	78.38	77.5	79	77.88	78.5	78.5	77.12	75.88
500MF4Gram	78.12	78.5	78	78.75	78.25	77.5	79.75	78	75.5
300MF4Gram	76.62	76.88	77.75	79.12	76.75	77.5	78.12	77.25	74.88
200MF4Gram	74.75	76	77.25	78.62	76.12	78.12	78.38	77.62	73.75
100MF4Gram	72.88	73.5	76.25	77	75.12	75.75	75.25	74.62	71.12
1000WekaWord	74.12	74.25	75.25	77	75.12	76.5	75.38	72.12	72.25
500WekaWord	75.75	75.88	74.38	77	74.75	76.88	73.5	72.75	74.62
300WekaWord	74.38	75.5	74.25	77.25	73.62	75.75	74.62	71.88	74.25
200WekaWord	73	72.88	73.75	77.62	72.88	73.88	73.25	68.75	71.75
100WekaWord	70.38	71	73.25	75.88	71.5	71.88	71.12	67.62	70.62
859Weka4Gram	77.12	77.25	78.12	79.25	76.75	78.38	77.88	77.38	74.62
500Weka4Gram	76.25	76.75	77.62	78.5	76.5	78.75	77.75	77.62	74.88
300Weka4Gram	75.5	77.25	78.25	78.75	77.25	78.88	77.62	77.25	74.75
200Weka4Gram	75.5	75.75	77.75	78.62	76.88	78.38	77.38	78.38	75.62
100Weka4Gram	75.38	75	78.38	79.88	75.5	77.38	76	76.75	73.75
810MenexMFW	75.25	75.12	73.88	77.5	74.5	76.62	77.25	73.12	72.75
500MenexMFW	75.5	75.62	76.12	78.75	74.5	76.88	77.5	75.75	72.25
300MenexMFW	75.75	75.62	76.88	77.88	75.5	76.75	77.75	76.25	73.5
200MenexMFW	75.38	75.62	77.75	78.38	74.38	78	77.88	76.75	72.12
100MenexMFW	74.75	74.88	78.12	78.75	75.5	78	77.62	76.38	71.62

For more detailed analysis we can construct a correlation matrix for the individual feature sets for all Platonic works. In the example given in FIGURE 2, which shows the correlation matrix for the most-frequent 4-grams, the exclusively pale peach row clearly shows that the *Menexenus* does not have a positive correlation with other Platonic works. The *Parmenides* (tlg009) is another dialogue that often has a weaker connection to the rest of the *Corpus Platonicum*. Additionally, we can see that there are Platonic works that do correlate. For instance, the *Sophista*, *Politicus*, *Philebus*, *Leges*, *Epinomis*, and *Letters* (tlg007, tlg008, tlg0010, and tlg034–tlg036) show a strong correlation (dark blue on the matrix in FIGURE 2). They are also all works that are traditionally considered to be late Platonic dialogues. Similarly, the *Clitopho* (tlg029) shows a correlation with the late Platonic dialogues, although when using other feature sets it sometimes shows no correlation or a very weak correlation with the *Menexenus*. The *Timaeus* and *Critias* (tlg031 and tlg032) always correlate strongly. Here they have a weak correlation with the late dialogues too.
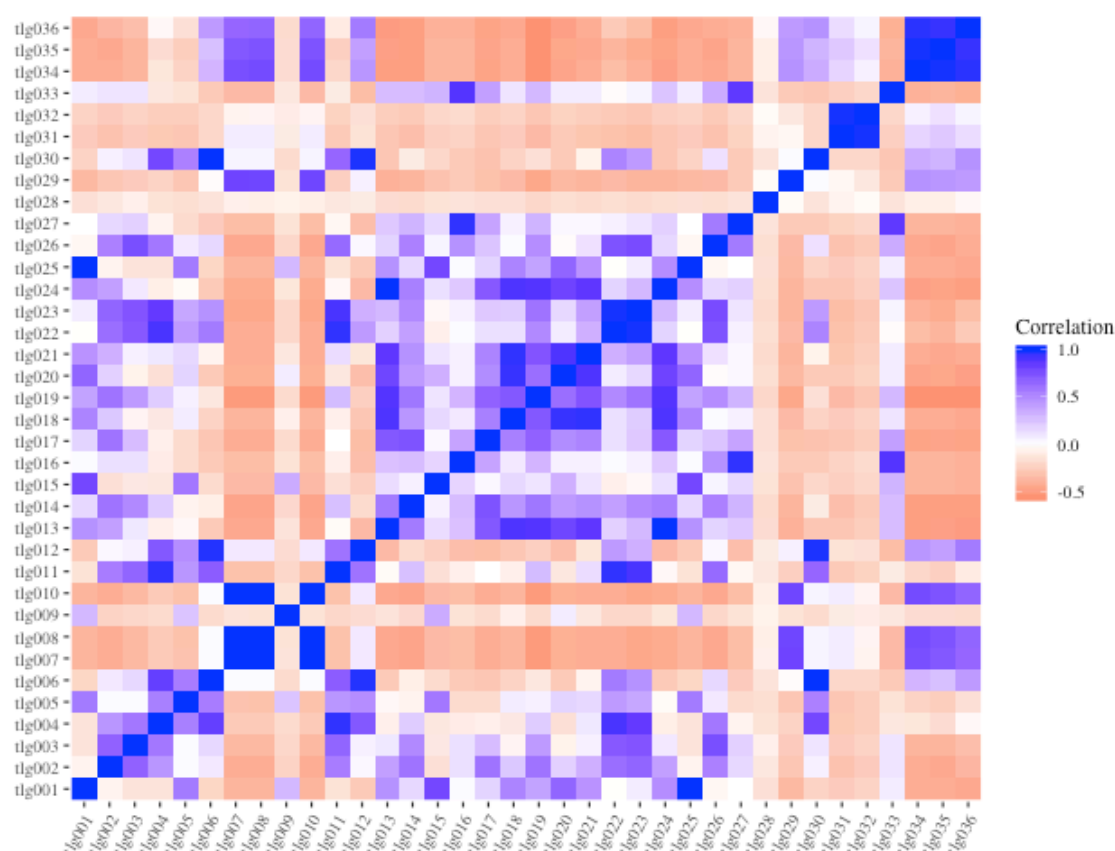
FIGURE 2: Correlation matrix for the *Corpus Platonicum* analyzed with the MF4-gram feature sets. The axes show the CTS Work Identifier (e.g., tlg028 is the *Menexenus*). We can see, for instance, that tlg034, tlg035, and tlg036 have a strong correlation with each other.

We can also use the correlation with other Platonic texts to rank all works of the *Corpus Platonicum* according to their 'Platonic-ness' of style. FIGURE 3 thus shows that the most 'Platonic' work seems to be the *Theaetetus* (tlg006); however, when I used other feature sets *Laches* (tlg019), *Euthydemus* (tlg021), and *Gorgias* (tlg023) also topped the group. While there was variance among the most 'Platonic,' every feature set ranked the *Menexenus* last, with very low correlation to other Platonic works. In fact, the *Menexenus* clusters better with Xenophon, albeit also not convincingly (see FIGURE 4), or even Lucian, which is impossible, because the *Menexenus* was definitely written before Lucian was born.[27]

---

[27] More visualizations of the data supporting this are included on the GitHub repository accompanying this article (the link has been changed for peer review):
https://drive.google.com/drive/folders/1wvmMNndl6utRCVqNvcPbtSxqdp-i_om-?usp=sharing .
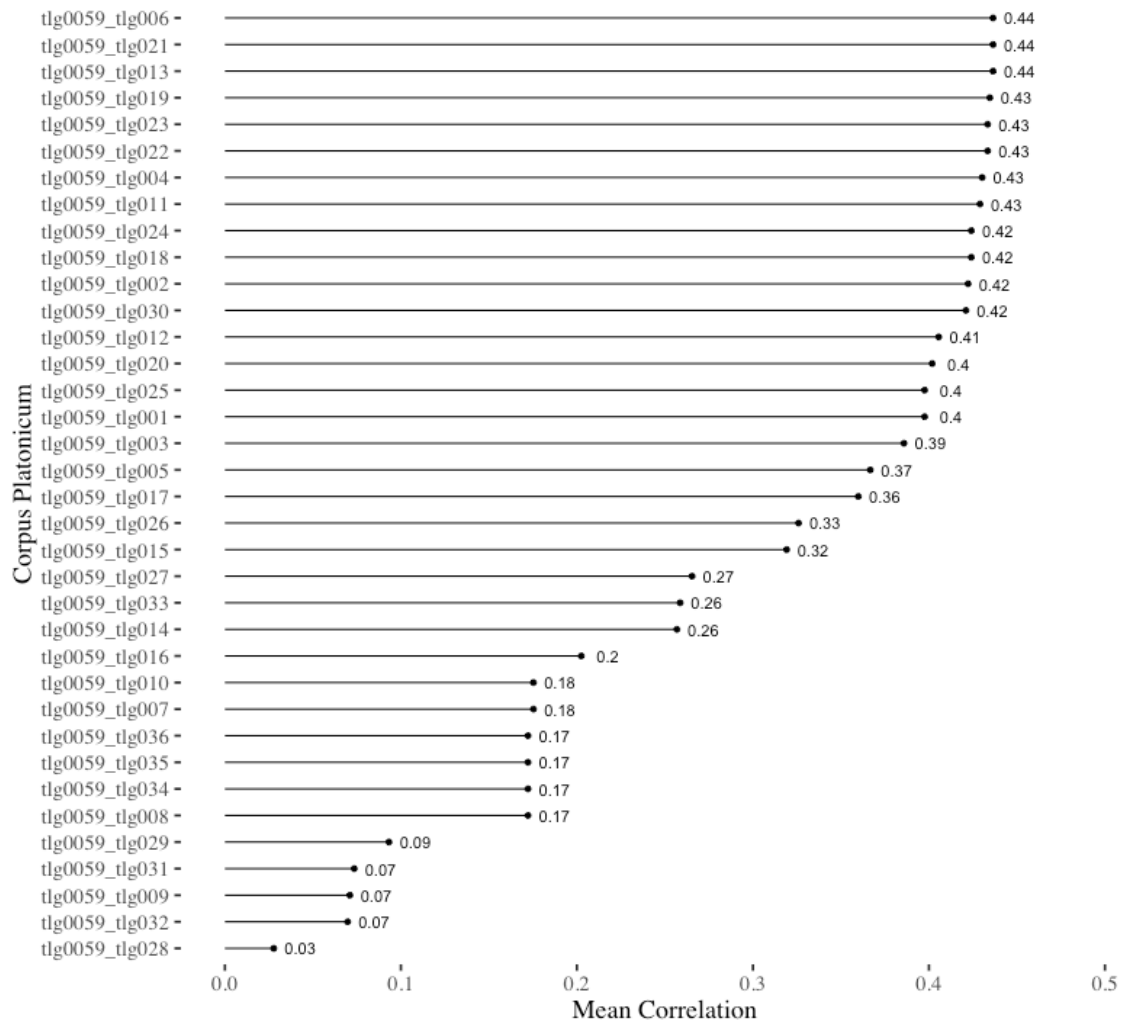
FIGURE 3: Works of the *Corpus Platonicum* analyzed with the MF4-gram feature sets and ranked by their correlation with other works
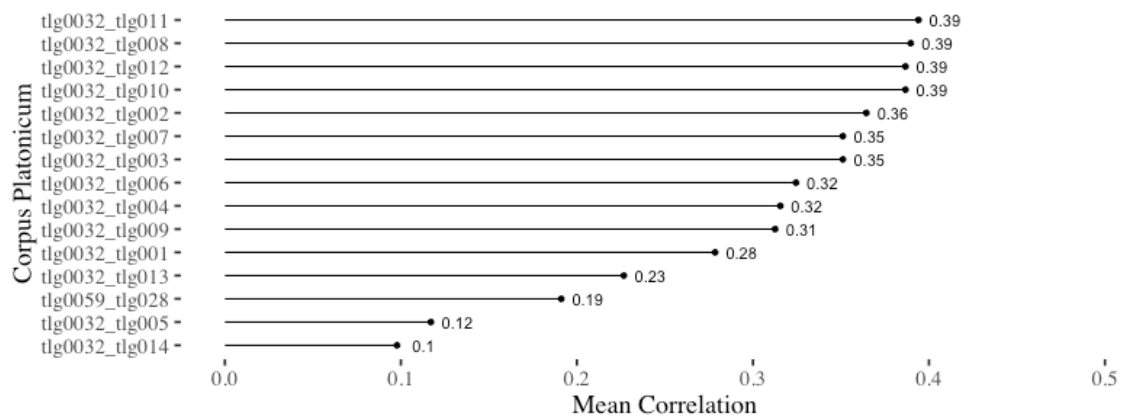

FIGURE 4: Works by Xenophon and the *Menexenus* (tlg0059_tlg028) analyzed with the MF4-gram feature sets and ranked by their correlation with other works

While those results are very conclusive and while many weak discriminators perform better than a few strong ones in authorship attribution tasks,[28] I was also interested whether I could find example words where the *Menexenus* is a clear outlier compared with the rest of the *Corpus Platonicum* to make the results more tenable for philologists working qualitatively. Outliers in a data set are calculatable values that are distant from the other observations. In statistics, every value that is 1.5 times outside the interquartile range below the lower or above the upper quartile can be considered an outlier.[29] When looking at the 5000 most-frequent-words table generated by *stylo*, one can identify 294 clear outliers (345 for the most frequent 4-grams). To only show words that are very common in the *Corpus Platonicum*, the number of outliers was reduced to words that occur at least once every 500 words (see FIGURE 5).
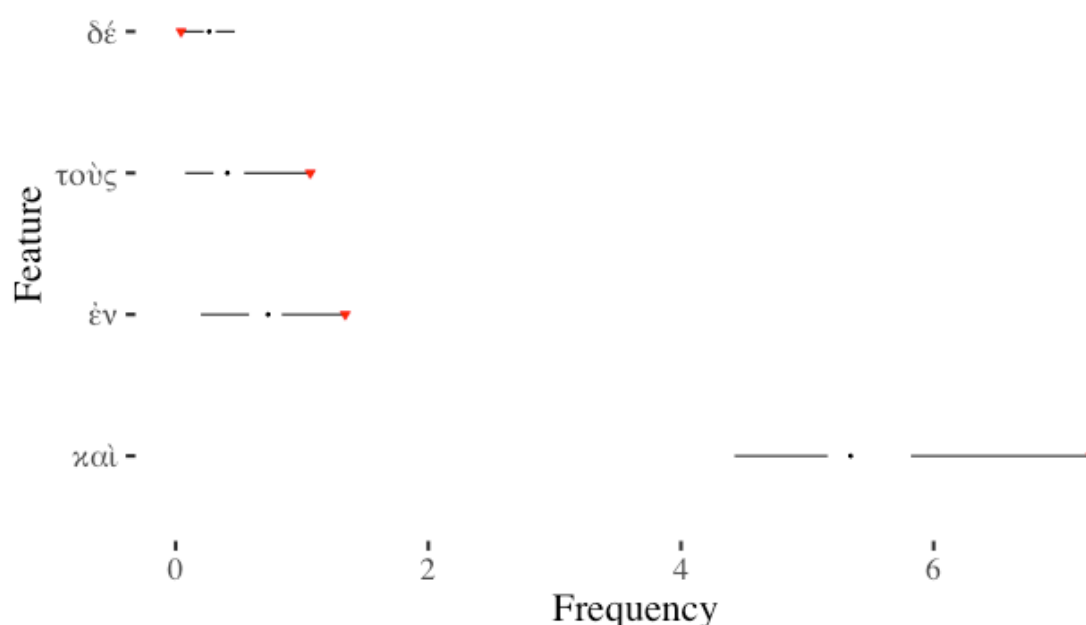

FIGURE 5: Most frequent words in the *Corpus Platonicum* where the *Menexenus* is an outlier and that occur at least once every five hundred words. The frequency is given in occurrence per one hundred words. The red triangle shows the value for the Menexenus.

Looking at FIGURE 4, the fact that 'καὶ' is a lot more frequent and 'δέ' is a lot less frequent in the *Menexenus* than in the *Corpus Platonicum* as a whole shows a different preference for connecting the parts of sentences that are on the same level. For 'καὶ' it almost does not matter whether we look at just the speech in the *Menexenus* or only the preceding and concluding dialogue, because even in the dialogue the frequency of 'καὶ,' at 6.42%, is on the upper limit of what has been observed throughout the *Corpus Platonicum*. Furthermore, if we

---

[28] J. Burrows, "Delta" 268.
[29] Calculated with the R function *boxplot.stats*.

just look at the character Socrates' uses of 'καὶ' throughout the *Corpus Platonicum*, then the *Menexenus* is a clear outlier too (see FIGURE 6). This claim becomes even stronger if we follow Sanson's argument that Plato changes his usage of conjunctions like 'καὶ' as specific linguistic markers to enhance the portrayal of his characters.[30]
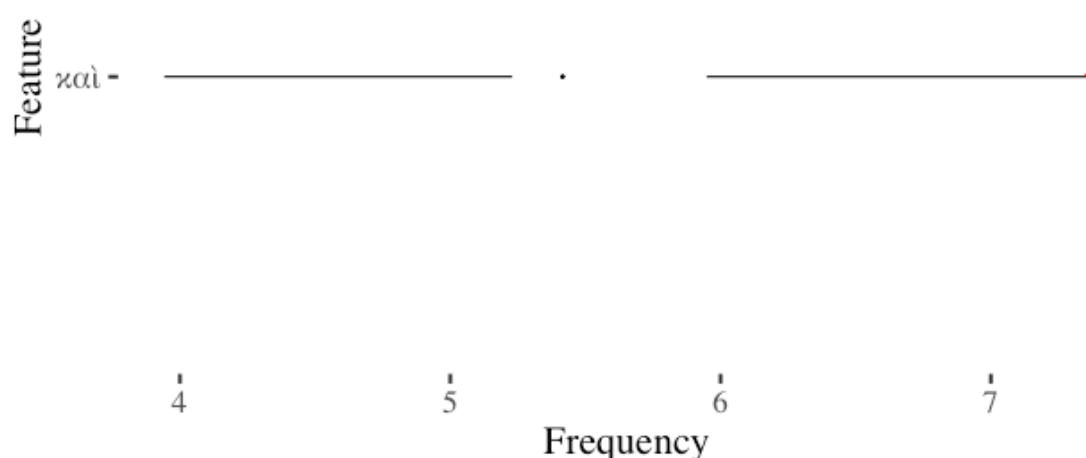

FIGURE 6: Use of 'καὶ' by the character Socrates in the *Corpus Platonicum*. The frequency is given in occurrence per one hundred words. The red triangle shows the value for the Menexenus.

Ledger suggested that such differences might be explained by a genre signal.[31] Although this sounds tempting, this hypothesis can be dismissed because, first, in the t-SNE visualizations orators like Isocrates, Lysias, Libanius, and Demosthenes have the tendency to cluster in close proximity, but the *Menexenus* does not. Second, when I repeated the co-occurrence experiment by pairing the *Menexenus* with the orators, it did not show a positive co-occurrence, and when I repeated the outlier experiment for the orators and the *Menexenus*, one sees that the use of 'καὶ' and 'ἐν' is as unusual in that context as it is when compared with the works of the *Corpus Platonicum* (see FIGURE 7). Based on this large-scale analysis Ledger's assumption can now be shown to be false.

---

[30] Sansone, "Stylistic Characterization" 173.
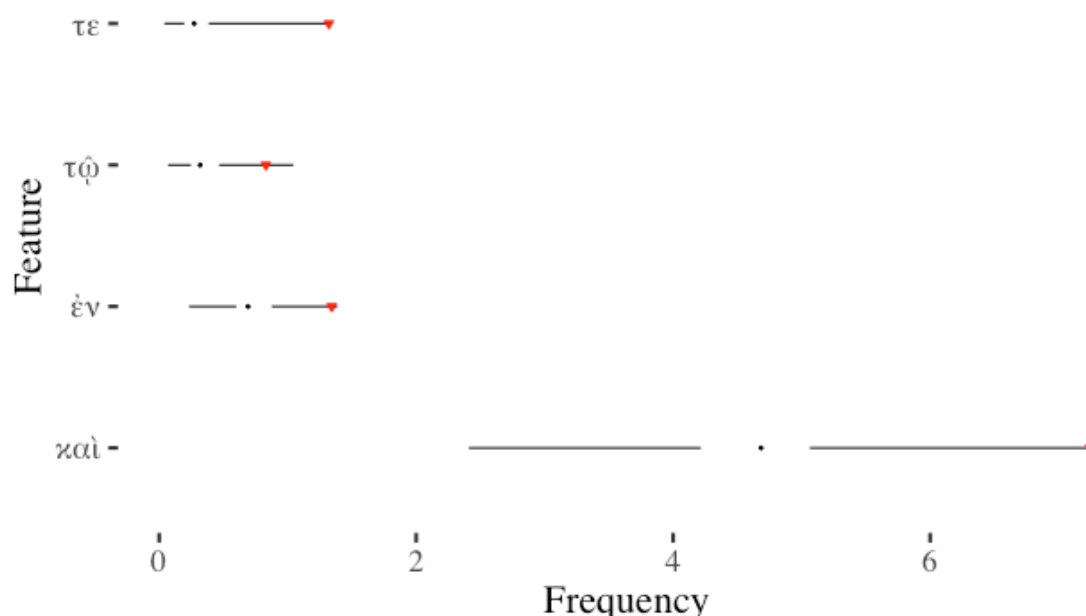[31] Ledger, *Recounting Plato* 163.

FIGURE 7: Most-frequent words used by the orators and the *Menexenus* where the latter is an outlier and that occur at least once every 500 words. The frequency is given in occurrence per one hundred words. The red triangle shows the value for the Menexenus.

All these stylometric findings—that the measures are reliable and mostly do not cluster the *Menexenus* with other works of the *Corpus Platonicum*, that the *Menexenus* has, in fact, a higher cooccurrence with Xenophon and Lucian, that its use of 'καὶ' among other examples is unusual, not only for other works in the *Corpus Platonicum* but also for how Plato's Socrates speaks— indicate strongly that the *Menexenus* was not written by Plato. Therefore, if the philological reasons for a Platonic authorship of the *Menexenus* are not beyond reasonable doubt, one can safely assume that the work was written by someone else.

The Philological Argument for and against Attributing the Menexenus to Plato

As mentioned above, for most of the nineteenth century the *Menexenus* was not thought to have been written by Plato. This flipped when influential philologists, such as Ulrich von Wilamowitz-Moellendorff, changed their opinion and considered the work Platonic.[32] All philological defenses of the hypothesis that the *Menexenus* was written by Plato broadly follow the same pattern, with one of the best summaries given by Stavros Tsitsirides.[33] Essentially dividing the arguments into three categories, Tsitsirides noted that the first is the ancient works

---

[32] See, for example, U. v. Wilamowitz-Moellendorff, *Platon: Beilagen und Textkritik* (Berlin 1920) 126 n.2.
[33] S. Tsitsirides, *Platons Menexenos: Einleitung, Text und Kommentar*, Beiträge zur Altertumskunde, Band 107 (Stuttgart and Leipzig 1998) 34–41.

that attribute the *Menexenus* to Plato, the second is the stylistic criteria, and the third category is the textual connections to other Platonic works.

The evidence for the first category can be reduced to two passages by Aristotle that refer to the *Menexenus*, since all other evidence mentioned by Tsitsirides dates to at least three hundred years later. One of the first references in the latter group (Cic. Or. 151), clearly attributes a eulogy to Plato, which was—according to Cicero, the only witness I could find for this statement—restaged by the Athenians every year. Given that such a well-received writer as Cicero attributes the *Menexenus* to Plato, it is also not surprising that from then on it was often taken as fact that the *Menexenus* was Platonic. That being said, Cicero and his contemporaries mark a *terminus-post-quem* for the *Menexenus* being publicly accepted as a Platonic dialogue. From that period onwards, it almost doesn't matter whether it was written by the authorial entity Plato or not, because even if it were not written by Plato, it was received as Platonic thought and influenced philosophical and rhetorical discussions accordingly.

While this is certainly important when studying the reception of Platonic thought, it does not directly support the hypothesis that the *Menexenus* was written by Plato, only that it was perceived as having been written by Plato from at least the first century BCE onwards. That said, it also wasn't always perceived in this way either: there is another ancient witness that makes a strong case against Platonic authorship. The late antique Diogenes Laertius wrote in his entry about Glaucon (Diog. Laert. 2.14): "Γλαύκων Ἀθηναῖος: καὶ τούτου φέρονται ἐν ἑνὶ βιβλίῳ διάλογοι ἐννέα: … Μενέξενος." Glaucon, Plato's brother, seems to have written nine works that were published in one book, of which Diogenes Laertius knew. Given that the stylometric analysis at least points (with the similarity to Xenophon) to an author who lived in a linguistically similar time and space as Plato, Glaucon would be a fitting alternate hypothesis for authorship of the *Menexenus*.[34] The fact that Xenophon, in his invective criticism of Glaucon in *Mem.* 3.6.1, describes how Glaucon had a vested interest in oratory and was training to become a politician makes the case even stronger. Glaucon's character and his having trained to become a public speaker seems a better fit with a fictional eulogy than with Plato's widely known criticism of such exercises. Unfortunately, Diogenes Laertius' statement is late and he also does not quote any of the text of the *Menexenus*, just its name. But we must ask, how likely is it that there was another work written by Plato's brother that was also called *Menexenus*? Or was Plato's *Menexenus* 'wrongly' attributed to Glaucon?

---

[34] The possibility that the author is Glaucon had already been suggested by F. Ueberweg, *Untersuchungen über Echtheit und Zeitfolge platonischer Schriften und über die Hauptmomente aus Platons Leben* (Wien 1861) 146.

The more we look into the younger witnesses the more it seems unclear who wrote the *Menexenus*. The only strong witness for it being written by Plato is Aristotle and it seems that the whole weight of the first category of arguments for a Platonic authorship can be reduced to whether or not we think that Aristotle is referring to Plato. Because of its importance, I will analyze the Aristotelean passages put forward by Tsitsirides as evidence after I have discussed the other two categories of arguments in favor of attributing the authorship of the *Menexenus* to Plato. Yet, I hope to have shown that other ancient witnesses, although they help us to understand the reception of Platonic thought, contribute little to answering the question of authorship.

Tsitsirides' second category is the stylistic evidence and it is very much a weak point of his defense of the Plato-attribution. While he suggests that it is possible that a person could have written the dialogue shortly after Plato's death, maybe even a person belonging to Plato's school, and that the work entered the transmission lines of the *Corpus Platonicum* by accident,[35] he dismisses evidence that supports this interpretation. He only glosses over Ledger's computational analysis, because he seems to have misunderstood its result, and focusses mainly on Ritter's nineteenth-century manual labor.[36] But Ritter's work uses only forty-three features of which only ten occur in the *Menexenus*; most of them only once.[37] Tsitsirides misconstrues Ledger's remark that he "concede[s] a weakness in the stylometric evidence" as that he concedes a weakness in the stylometric *method*. It seems clear that Ledger meant "weakness in the stylometric evidence" that the *Menexenus* was written by Plato. He clearly states that the "verdict is rather against" the *Menexenus* being written by Plato, but he is "inclined to think that that is more a reflection of its peculiar nature than a definitive declaration against Platonic authorship."[38] Thus, while Ledger considers the *Menexenus* a Platonic work, he does so against the statistical evidence, which he dismisses through his assumption that this might be entirely based on a genre signal.[39] Although Ledger explains that the genre signal does not seem to influence, as one might expect, the letters, including the seventh letter, he is so convinced by the non-stylometric evidence provided by the Aristotelean references that he does not even include the *Menexenus* in his discriminant analysis authenticity study.[40]

---

[35] S.Tsitsirides, *Platons Menexenos* 37.
[36] S.Tsitsirides, *Platons Menexenos* 38–39.
[37] C. Ritter, *Untersuchungen* 99.
[38] Ledger, *Recounting Plato* 163. He also mentions (169) that although the stylistic evidence is against them, he thinks that both dialogues, *Menexenus* and *Clitopho*, seem "too outrageous to have been written by anyone other than Plato."
[39] Ledger, *Recounting Plato* 163.
[40] Ledger, *Recounting Plato* 105, 163.

Given, as shown by the stylometric analysis above, that the *Menexenus* looks less likely to have been written in Platonic style the more data we include in our analysis, it becomes obvious that Tsitsirides' stylistic arguments are partially cherry-picking, partially based on Ledger's doubting his own results, and partially simply ignoring evidence. The history of twentieth-century research on the question of whether or not the *Menexenus* was written by Plato was mainly driven by researchers who really *wanted* it to be written by Plato. An unbiased look at Tsitsirides' second category of arguments, instead, finds more evidence against a Platonic authorship of the *Menexenus*.

The third category of arguments is the similarities in content of the *Menexenus* and other Platonic works. Tsitsirides admits that the lack of clear philosophical content in the *Menexenus* makes it difficult to compare it to Platonic thought evident in the other dialogues. Thus, Tsitsirides resorts to listing individual intertextual connections that in their sum clearly show a familiarity with Platonic works by the author of the *Menexenus*. Aside from one passage, Tsitsirides does not attest a direction for the intertextual connections and we do not know which work came first. The only passage where he sees a direction is *Menex.* 240a–b and *Leg.* III 698 c–e. There, Tsistirides follows Pohlenz's argument, which essentially states that the *Menexenus* must have been earlier because it is more detailed.[41] I fail to see the direction of the intertextual reference as proven, especially given that the story also features in Herodotus (Hdt. *VI*.31). That being said, the third category of arguments certainly proves familiarity with the other Platonic works by the author of the *Menexenus*. Given that the stylistic evidence speaks against the *Menexenus* being written by Plato and that the intertextual evidence at the very least shows a connection between the *Menexenus* and other Platonic works, and given that later ancient sources might be influenced by their reading of Aristotle and Cicero, the complexity of the question of whether the *Menexenus* was written by Plato can be entirely reduced to whether or not Aristotle attributes the *Menexenus* to Plato. If he does, then the many stylistic anomalies of the *Menexenus* might be just a sign of Plato's wide stylistic range. If Aristotle does not conclusively attribute the *Menexenus* to Plato, then the stylistic evidence outweighs the testimony of later ancient authors and individual intertextual connections to the *Corpus Platonicum* and the *Menexenus* must be considered not to have been written by Plato.

---

[41] Tsitsirides, *Platons Menexenos* 41. See also M. Pohlenz, *Aus Platos Werdezeit: Philologische Untersuchungen* (Berlin 1913) 278–281.

Aristotle and the Menexenus

In his *Rhetoric*, Aristotle does not name Plato as the author of the *Menexenus*, nor does he even name the *Menexenus* itself. He does, however, refer to the same passage from the *Menexenus* (235d) in two instances. The first instance is *Rh*. 1.9.30:

> σκοπεῖν δὲ καὶ παρ᾽ οἷς ὁ ἔπαινος· ὥσπερ γὰρ ὁ Σωκράτης ἔλεγεν, οὐ χαλεπὸν Ἀθηναίους ἐν Ἀθηναίοις ἐπαινεῖν.

> but let's also look at with whom present we praise: for as Socrates said, it is not difficult to praise Athenians among Athenians.[42]

It has been argued and accepted that the use of the imperfect ἔλεγεν by Aristotle makes it in this instance more likely that he references the historic Socrates.[43] Given the almost proverbial nature of the passage, it is clear why this was accepted. The passage in the *Menexenus* (235d), however, is much more detailed and in fact slightly inverse. While Aristotle simply states that which is not difficult, the Socrates in the *Menexenus* emphasizes how difficult it would be to praise someone in the presence of their enemies:

> εἰ μὲν γὰρ δέοι Ἀθηναίους ἐν Πελοποννησίοις εὖ λέγειν ἢ Πελοποννησίους ἐν Ἀθηναίοις, ἀγαθοῦ ἂν ῥήτορος δέοι τοῦ πείσοντος καὶ εὐδοκιμήσοντος· ὅταν δέ τις ἐν τούτοις ἀγωνίζηται οὕσπερ καὶ ἐπαινεῖ, οὐδὲν μέγα δοκεῖν εὖ λέγειν.

> For should there be the need of speaking well of Athenians among Peloponnesians or Peloponnesians among Athenians, then it would need a convincing and highly esteemed speaker: but whenever someone contests in the presence of whom they are praising, speaking well doesn't seem to be a big thing.

One can see that the first Aristotelean passage alone does not even prove the existence of a dialogue let alone the authorship. It could simply be an anecdote that was told about Socrates. The second Aristotelean instance (*Rh*. 3.14.11), however, is not so easy to dismiss:

> ὃ γὰρ λέγει Σωκράτης ἐν τῷ ἐπιταφίῳ, ἀληθές, ὅτι οὐ χαλεπὸν Ἀθηναίους ἐν Ἀθηναίοις ἐπαινεῖν, ἀλλ᾽ ἐν Λακεδαιμονίοις.

> For Socrates says in his funeral oration correctly, that it is not difficult to praise Athenians among Athenians, but among Lacedaemonians.

I will only mention in passing, because I think it has been sufficiently disproven, that critics have considered the third book of the *Rhetoric* itself where Aristotle refers to the *Menexenus* as unauthentic. After Diels showed that the third book of the *Rhetoric* is Aristotelean, however,

---

[42] All translations are my own.
[43] Tsitsirides, *Platons Menexenos* 34; Ueberweg, *Untersuchungen* 143.

people then assumed that the passage was an interpolation.[44] I agree with Tsitsirides that this is unlikely: not least because—aside from everything else—the choice of the word ἐπιταφίῳ seems odd for someone who interpolates a clarification.

Thus, although an interpolation would tip the scales entirely to assuming that the *Menexenus* was not written by Plato, let's for the sake of argument assume the more likely scenario that the reference is authentically Aristotelian: despite its brevity, reverse order, and Aristotle reducing the Peloponnesians to the Lacedaemonians, the Aristotelean passage has all elements of the *Menexenus*. The present of λέγει and the naming of the work ἐν τῷ ἐπιταφίῳ also seem to point to a written resource and not necessarily to the historic Socrates.[45] Interestingly though, Aristotle calls the work *Epitaphios* and not *Menexenus* and he doesn't name Plato as the author. That is probably also the reason why Bonitz, although he lists the passages under the works by Platon in his *Index Aristotelicus*, refrains from categorizing them under group 'a,' the group he reserves for references where either Plato or Socrates are mentioned and a Platonic work named.[46] Instead, Bonitz refers to Ueberweg's theory that Glaucon is the author,[47] so the reader can judge for themselves whether this is a clear Plato reference or not. Ueberweg points out that although the formula ὃ γὰρ λέγει Σωκράτης ἐν τῷ ἐπιταφίῳ fits with how Aristotle, at times, refers to Platonic works, it does not mean that this formula absolutely must refer to a Platonic work.[48] In order to judge whether Aristotle refers to Plato or not we must thus look at all references to Plato or Socrates in Aristotle's *Rhetoric*.

In his *Rhetoric*, Aristotle refers to Plato by name only three times. In the first instance (1.15, 1376a) it is not entirely clear whether he means the philosopher or the comic writer Plato. Whomever he means, it is the historical person not the work of an author. In the second instance, Aristotle seems to be critical of how Socrates is portrayed by Plato when he reports in 2.23, 1398b that Aristippus corrects Plato saying that Socrates never spoke like Plato made him. It is interesting that Aristotle's unveiled criticism of Plato's dogmatic portrayal of Socrates is in the same work where it is not clear whether a statement by Socrates is by Plato or not. With that in mind, one would have hoped that Aristotle would have referred clearly to Plato when mentioning Plato's Socrates, albeit that he often means the Platonic Socrates when speaking of

---

[44] Tsitsirides, *Platons Menexenos* 35 and H. Diels, *Über das dritte Buch der aristotelischen Rhetorik* (Berlin 1886) esp. 20–23. Diels thinks that the *Menexenus* might not be Plato's masterpiece (in fact he thinks it is a "taube Nuß") but he considered it authentic, because he considers the third book of Aristotle as authentic.
[45] Ueberweg, *Untersuchungen* 144.
[46] H. Bonitz, *Index Aristotelicus* (Berlin 1870), 598, s.v. Πλάτων.
[47] Ueberweg, *Untersuchungen* 146.
[48] Ueberweg, *Untersuchungen* 144.

just Socrates. But only in the third instance, is Aristotle clearly referring to Platonic passages (3.4, 1406a).

Aside from the two passages in question, Aristotle refers to a Socrates ten times. The first passage (1.1, 1356b) has the name Socrates twice and then only as a stand-in for a named individual or an individual case, while in the second passage (1.2, 1357b), a statement about the historic Socrates is given as an example to clarify a point. The third instance, again, only refers to Socrates as a named individual (2.4, 1382a). The next passage (2.15, 1390b) again refers to the historic Socrates, or rather his descendants, as an example. In the fifth passage (2.20, 1393b), Aristotle describes the effect of Socratic comparisons. If this points at all to a written account, then Xenophon's *Memorabilia* have the closest parallel in 1.2.9. Given the context (that Xenophon reports on one of Socrates' accusers), it is much more likely that Aristotle refers to the historic Socrates. Similarly, in 2.23, 1398a, Aristotle refers either to Xenophon (*Apol. Socr.* 17) or the historic Socrates. Passage 2.23, 1399b is the criticism against Plato's Socrates and 2.23, 1399a refers to Theodectes' Socrates. If 2.23, 1399b is Socrates and not Isocrates, Aristotle is then referring to Socrates as an example. In the last instance (3.18, 1418b), Aristotle seems to point to Plato's *Apology* 27c–d or to the historic Socrates. If he points to Plato, then the way he paraphrases Plato's work indicates that he referenced Plato by heart rather than from the dialogues in front of him. The name Plato is indeed missing, which would support a reading that the passage 3.14, 1415b at least points to Plato as the author of the *Menexenus*.

That being said, in his *Rhetoric*, Aristotle refers to Socrates twelve times: three times he may be referring to Plato without mentioning him by name (if we include the two passages to the *Menexenus*, that is); twice he refers to Xenophon, also without mentioning his name; once he mentions Theodectes as the author; four times he mentions Socrates as a historical example; and, finally, twice he just uses him as a variable for any individual person. Of the six passages that might refer to another work, the author of the work is only mentioned once. The other five passages have in common that they seem like stories known to everybody at that time and Aristotle is barely using the same words as the other authors: he paraphrases in all five instances. We do not know whether Aristotle's information comes from reading those works or from listening to those well-known stories; stories that in turn seem to tell us more about the historic Socrates than that Aristotle is unquestionably pointing to Plato. After all, he also mentions stories that occur in Xenophon, without mentioning Xenophon by name. That means in regard to the Aristotelean *Menexenus* reference in question, as Ueberweg noted, just because Aristotle sometimes seems to refer to a Platonic work by only mentioning Socrates and the

work (and not Plato himself), it doesn't mean that he must be doing so in this very instance.[49] Especially, when it is something so entirely proverbial as the observation that it is not difficult to praise Athenians among Athenians.

Conclusion

This paper might actually praise Lacedemonians among Athenians when it deviates from the generally accepted hypothesis that the *Menexenus* was written by Plato. Yet, as Ledger has stated and I have now shown more thoroughly above, there is very little stylistic evidence that the *Menexenus* was written by Plato. The stylometric measures have been tested and seem reliable, and mostly do not cluster the *Menexenus* with other works of the *Corpus Platonicum*. The *Menexenus* has a higher cooccurrence with Xenophon and Lucian than with platonic works, while the frequency of 'καὶ' in the *Menexenus* is unusual not only for other works in the *Corpus Platonicum*, but also for how Plato's Socrates speaks. Furthermore, the genre signal hypothesis could be disproven. This together indicates strongly that the *Menexenus* was not written by Plato.

The philological evidence, on the other hand, tells us that the author certainly shows great familiarity with the *Corpus Platonicum*, as Tsitsirides has shown, and the stylistic evidence also shows similarities with other Athenian authors from the fifth and fourth century BCE. Ancient testimonies are too late to be useful, aside from Aristotle, and yet if one looks at how Aristotle references the *Menexenus*, it is not clear whether he references the *Menexenus* at all, whether he references Plato's *Menexenus*, or whether he just references a story about the historic Socrates. It seems tempting to attribute the *Menexenus* to Glaucon, as Diogenes Laertius may have done and as Ueberweg suggested. But there is of course the fact that the *Menexenus* was transmitted in the Platonic canon of the tetralogies, which seems to suggest that the burden of proof lies heavier on the side arguing against Platonic authorship.

Depending how much weight the reader puts on the stylistic evidence or on the fact that the second of the two Aristotelean passages is similar to passages where Aristotle refers to Platonic works by just mentioning Socrates and not Plato, the answer to the question whether or not the *Menexenus* was written by Plato seems to change. Yet, the fact that might tip the scales, because it qualifies the argument that Aristotle must have meant the Platonic Socrates, is the fact that Aristotle also refers to a non-Platonic Socrates frequently in his *Rhetoric*. With all this in mind a reader must at least consider the possibility that the *Menexenus* was not written by Plato. The stylistic evidence, at least, now strongly indicates that it wasn't.

---

[49] Ueberweg, *Untersuchungen* 146.