Pythagoras said that the nature of all things is number (fr. 58 B 4,8 DK⁶ = Arist. *Metaph.* 1.5 [985b23-6a15, 987a9-10]), but Aristotle cautions that we must not expect more precision than the subject warrants (*NE* 1.3.1 [1094b12-4]). It is between these pillars that stylometry has fallen; Brandwood hopes to set it right, for Plato at least, by dint of greater breadth. Alas, he founders on a methodological rock, rarely perceived.

Although there is some evidence that the notions of measurement uncertainty (Thuc. 3.23, e.g.) and of combinatorics (Xenokrates *apud* Plut. *QC* 733a, *Stoic Repugn.* 1047c) had crossed the minds of Greek thinkers,¹ and although the assigning of numerical values to letters and words (isopsephy) was a standard feature of Greek thought,² it was apparently left for mid-XVIII cent. A.D. Swedish theological debates to prompt the analytic application of number to texts, first by a clergyman named Kumblaeus.³ With the development some decades later of modern statistical methods (by Gauss in 1801-9 in connection with the orbit of an asteroid⁴), the way was opened for stylometry (though that name was coined much later). The theologian Friedrich Schleiermacher (1768-1834) in 1807 seems to have been the first to attempt the way, with the incidence of 'hapax legomena' in [Paul] *I Timothy*.⁵ Beginning in 1867 with the work of Lewis Campbell (see below) stylometry was applied to Plato's works, mostly with a view to establishing their chronology (for a few works, most infamously *Letter 7*, authorship is a question).

Brandwood's book is a revision of his 1958 London U. doctoral dissertation, "The dating of Plato's works by the stylistic method," apparently delayed by the production of his great *Word Index to Plato* (1976), and his recovery from a brain haemorrhage (p. x). In fact the revisions are not great and some chapters


² Paul Perdrizet, "Isopsephie," *REG* 17 (1904) 350-360.


⁵ I use 'hapax legomenon/a' relatively, to refer to words which occur once in a given text or author. Fr. E. D. Schleiermacher, *Ueber den sogenannten ersten Brief des Paulos an den Timotheos* (Berlin 1807), and repr. in Friedrich Schleiermacher's *sämtliche Werke* 1.2 (Berlin 1836) 221-320; the relevant passage is pp. 27-76 (pp. 233-254 of the reprint); this is followed by a study of words common to *I Tim.* and to *II Tim.* or *Titus* (pp. 77-104 = 254-65).
reappear almost *uerbatim* (e.g., c. 2 on Campbell). Portions have been published elsewhere: the last chapter of the thesis (pp. 407-25), somewhat revised, appeared as "Analysing Plato's style with an Electronic Computer," *BICS* 3 (1956) 45-54, and, with further revisions, as c. 4 (pp. 50-65) of *Mechanical Resolution of Linguistic Problems*, Andrew D. Booth, L. Brandwood (still B.A.), and J. P. Cleave (New York/London 1958). On the other hand, some material has been added, notably cc. 21 and 22 (pp. 228-48), analysing studies which appeared after the thesis (p. ix).

The book takes the form of a brief introduction (2 pp.) on the external evidence and the *status quo ante*, followed by 21 cc. summarising various studies, arranged (in all but one case, c. 18) chronologically, and fitted with a brief conclusion (4 pp.) and a set of 4 brief indices (4 pp.). The lack of a bibliography is mitigated by the treatment of the 22 studies in the individual chapters. Because the book is essentially a review of previous work, it must be covered in detail to appreciate B.'s accomplishment. In many cases the impression a reader would gain (that there is evidence for some particular view of the chronology or other) is wrong, and it is important to take the opportunity to correct these.

Lewis Campbell, *The Sophistes and Politicus of Plato* (Oxford 1867) was concerned with the dates of the two dialogues, and made six observations tending to date them late, of which B. subjects two only (# 5—'rhythmical cadence' and # 6—'unusual words') to statistical treatment or discussion (pp. 3-8). Campbell only grouped the *Politicus* and *Sophistes* with other late works, and despite "errors of judgment" (p. 8) is praised for moderation in not overusing his figures. Campbell's figures for vocabulary—still difficult to analyse numerically, though B. fails to explain this (cp. below on Dillenberger and μην)—are computed as the number of words a given dialogue had in common exclusively with *Tim., Crit., and Laws* (considered as a set). Campbell conservatively noted the effect of subject matter and opportunity, and refrained from drawing detailed conclusions. As to rhythm, Campbell only noted and exemplified the 'peculiar, stately rhythm'—here B. checks Campbell more carefully and concludes "Campbell had a flair for recognizing important stylistic phenomena without drawing up anything like complete statistics for them" (p. 6).

Friedrich Blass, in *Die attische Beredsamkeit* (Leipzig 1874) 2,426, remarked that Plato later tended more to avoid hiatus, and supported this with figures which are computed as number of incidences of hiatus per Teubner-page. This gives an order: *Laws* I (29), *Phil.* (20), *Tim.* (6), *Soph.* (2), *Pol.* (1), where I have given the average number per ten Teubner-pages in parentheses. This is not the correct computation, for it takes no account of opportunity for hiatus (which will depend on the number of words in the text ending and beginning with vowels). B. sees the importance of Blass' work in its confirmation of Campbell's based on "entirely different methods" (well, at least a different measure). In fact Campbell's one numerical result gives for these works the order:

---

1The first three — Socrates not chief speaker, *Sophistes* and *Politicus* form middle of unfinished tetralogy to which structure compare *Timaeus* and *Critias*, and didactic tone — are indeed not statistical. But why not # 4 'word order'?
Phil., Soph., Pol., Tim., Crit., Laws (the last three assumed late and like), with others intervening—hardly "exactly the same conclusion" (pp. 9-10).

W. Dittenberger, "Sprachliche Kriterien für die Chronologie der platonischen Dialoge," Hermes 16 (1881) 321-45, investigated the occurrences of various uses of μήν (pp. 11-22). In particular τί μήν; (as an affirmative reply), γε μήν and ἀλλὰ...μήν appear only in Symp., Lys., Phdr., Rep., Parm., Phil., Soph., Pol., Laws, Tim., and Crit. (Ritter later noted that οὐ μήν appears also only in this group).¹ Ritter noted a difficulty (which inheres in any vocabulary-based measure)—in essence, with respect to what does one count? Dittenberger counted with respect to pages of text, but this does not take into account the opportunities (equivalently, the expected rate), and Ritter preferred to find the total number of occurrences of replies and determine what fraction of these were γε μήν or whichever (p. 17—and see below). Dittenberger tried to confirm his result with a study of the use of μήν in other Attic authors (pp. 13-4). Frederking (and Kugler) noted that the distribution of the various uses of μήν appeared to vary within works (p. 17). Dittenberger then examined Plato's use of comparative particles (διότερον and καθότερον), but the results are fuzzy at best (pp. 19-20). B. tries (p. 19) to make them more precise by noting a decreasing rate, counted as uses per page, from Euthd., Meno, Gorg., Crat., Phdr., Symp., Lysis (0.80) to Phdr., Rep., Theaet. (0.61), and thence to Parm., Phil., Soph., Pol., Tim., Crit., and Laws (0.40). In forming these groups B. has begged the question (he uses a presumed approximate chronological order), he has counted not with respect to total number of replies (cp. above), and in comparing the three numbers he claims a clear trend but he fails to determine whether it is in fact statistically significant. Briefly the uncertainty σ for each ratio would be respectively 0.80 ± 0.15, 0.61 ± 0.08, and 0.40 ± 0.07, and the level of significance of the pairwise differences would be z = 1.1 (73 %) and z = 2.0 (95 %).² That is, those are the probabilities that the difference is significant—and most statisticians would consider anything less than 95 % (I and most physicists prefer 99 %) as insignificant. B. calculates (pp. 21-2) various ratios, which suffer from the same defect. Still, it is fair to say that Dittenberger showed that something was going on as revealed by the relative rates of the use of certain particles.

A. Frederking, "Sprachliche Kriterien für die Chronologie der platonischen Dialoge," N Jbb f Phil 125 (1882) 534-41, attempted to limit Dittenberger's work by noting that particle use varies for many reasons, and does so within a work (pp. 237). That is, he was concerned about what a statistician would call the stability of the tests being used: a very important but usually neglected issue (cp. the Conclusion). He examined the use of γε μήν

¹To be slightly more precise: τί μήν; is absent from Symp., Tim., Crit., ἀλλὰ...μήν from Tim., Crit., γε μήν from Lys., and οὐ μήν from Symp., Phaedr., and Crit.

²The σ's are calculated from the data in B.'s Table 4.3 (p. 20), collecting respectively 7, 3 and 7 terms (one for each dialogue). For a description of σ and z and relevant bibliography see BMCR 2 (1991) 426.
within the *Laws*, and τί μήν; in the *Laws* and *Rep.* (above), as well as δή γε, unaccompanied τε (i.e., without corresponding καί, τε, or οὐ τε), μᾶν, and even ἐπὶ οὖ. When B. analyses earlier work carefully he generally includes helpful tables (such as for Dittenberger's μήν-statistics), but here he includes none, which obscures Frederking's point. But the point is stated clearly (p. 27): "the danger of making hasty deductions from insufficient material."

F. X. Kugler, *De particula τοι εἰσίων compositorum apud Platonem usu* (Diss. Basel 1886), showed in a similar way that the use of τοι varied considerably even within a work (pp. 28-33). Here again the problem of with respect to what to count arises. Kugler considered the ratio of μέντοι το τοίνυν (but both are absent from *Tim.* and *Crit.*, and οὖν appears for the latter), but (though B. does not say so) the numbers are too small to give statistically significant differences in many cases. If we calculate the variance and arrange the results chronologically (assuming that the *Laws* is late) we get:

<table>
<thead>
<tr>
<th>Text</th>
<th>Variance</th>
<th>Mean ± Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Phdo.</em></td>
<td>1.80</td>
<td>1.80 ± 0.50</td>
</tr>
<tr>
<td><em>Theat.</em></td>
<td>1.03</td>
<td>1.03 ± 0.23</td>
</tr>
<tr>
<td><em>Rep.</em></td>
<td>0.77</td>
<td>0.77 ± 0.10</td>
</tr>
<tr>
<td><em>Crat.</em></td>
<td>0.72</td>
<td>0.72 ± 0.20</td>
</tr>
<tr>
<td><em>Soph.</em></td>
<td>0.24</td>
<td>0.24 ± 0.07</td>
</tr>
<tr>
<td><em>Phil.</em></td>
<td>0.15</td>
<td>0.15 ± 0.06</td>
</tr>
<tr>
<td><em>Laws</em></td>
<td>0.14</td>
<td>0.14 ± 0.04</td>
</tr>
</tbody>
</table>

A trend is clear (and even significant, overall), but is it meaningful? (Kugler also counted the total number of τοι-compounds per Stephanus-page, less correctly as that does not take opportunity into account: cp. B., p. 32). Is it correct here to examine ratios rather than simple differences? And is it correct to consider these two τοι-compounds as effective synonyms? But Kugler's point, which B. unnecessarily disparages (it is more than "a Frederking-inspired joke at the expense of Dittenberger and all prospective 'stylometricians'"), is correct—"no one should draw such important conclusions as Dittenberger has done on the evidence of one or two particles" (p. 32).

Morris Schanz, "Zur Entwicklung des platonischen Stils," *Hermes* 21 (1886) 439-59, studied Plato's use of several synonyms to express 'really' or 'truly' (in contrast to 'seemingly' *uei sim.'): τοῦ δόντι, δόντως and ὠς ἀληθῶς, τῇ ἀληθείᾳ, ἀληθῶς, ἀληθείᾳ (pp. 34-40). Here the difficulty is again: with respect to what to count? To take a few clearer cases, *Gorg.* has 17 τῷ δόντι and no δόντως, *Phdr.* has 8 to 6, and *Phil.* has no τῷ δόντι and 15 δόντως. If we assume it is correct to take simple differences, the trend is quite significant. Given the *Laws* with 52 δόντως and no τῷ δόντι, it may be meaningful as well. One could add the *Gorg.* with 9 τῷ δόντι and no δόντως, the *Rep.* with 41 to 9, the *Soph.* with 1 to 22, and the *Phil.* and *Pol.* with none to 11 and 8 respectively (the figures for ἀληθῶ- are less clear). In summary, given the assumption the *Laws* are late, a probable order *Gorg.* and *Phdo.*; *Rep.; Phdr.*; *Tim.*, *Pol.*, *Phil.*, *Soph.*, and *Laws* (where the groups separated by semicolons could be defended statistically). Schanz himself attempted much finer distinctions which cannot be supported by the statistics, though B. discusses a number of them.

---

1 Should note that in this list no successive pair is significantly different, and some pairs (*Rep.* and *Crat.*; *Phil.* and *Laws*) are essentially identical.
E. Walbe, *Syntaxis Platonicae Specimen* (Diss. Bonn 1888), investigated πας and compounds (pp. 41-7). He found that Plato's use of σοματος/συνάπας increased over time (a much larger number in *Soph.*, *Pol.*, *Phil.*, *Tim.*, *Crit.*, and *Laws*), but that ἂπας was irregular. B. notes (p. 44) that various other factors (subject, hiatus, rhythm) play a role. B. following Lutosławski calculates the number of πας, etc. per standard page of text (which as noted above is not correct), and does so also for individual books of *Laws* and *Rep.* Although B. does not explore this, it is relatively clear that, in so far as such a number is accurate, the data can be naturally grouped into two sets (all the others versus the six dialogues listed just above) with average values of 2.16 ± 0.37 uses of πας per Stephanus page and 4.14 ± 0.47 per Stephanus page respectively (giving a z = 3.3, or P = 99.9 % probability of significance). The variation between books of the *Rep.* and between books of the *Laws* is large but stays within the limits expected based on the above averages: 2.12 ± 0.49 for *Rep.* and 3.97 ± 0.86 for *Laws.* That the variation should be so much larger (given that the *Rep.* and *Laws* are the two largest works and so should have the smallest merely statistical variation) suggests either that both works were written over a long period (as Siebeck, Janell and others have suggested, at least for *Rep.* I—see below—which indeed is rather lower), or that the measure is not a stable one. B. correctly notes that one can only draw 'the broadest conclusions' from these data (p. 44), but then tries to canvass various details (pp. 46-7).

H. Siebeck, *Untersuchungen zur Philosophie der Griechen* (Halle 1888), in an appendix examined Plato's use of question and reply formulae (pp. 48-54). B. found that Siebeck's figures for the number of simple direct questions introduced with ἄρα were often wrong (total number of simple direct questions was not checked); but he does not note that the percentage gives an order rather different than B.'s table would suggest (the table is B.'s arrangement—p. 49, n. 2). We find (with B.'s corrected percentages, and omitting a few dialogues for which the percentage would have a large σ):

- *Meno* = 11 ± 3  *Gorg.* = 13 ± 2  *Euthd.* = 16 ± 3  *Theaet.* = 18 ± 3
- *Prot.* = 19 ± 4  *Rep.* = 21 ± 2  *Phdo.* = 22 ± 4  *Crat.* = 24 ± 4
- *Lysis* = 25 ± 6  *Soph.* = 29 ± 5  *Parm.* = 29 ± 4  *Laws* = 34 ± 4
- *Phil.* = 35 ± 5  *Pol.* = 36 ± 7

As usual B. has failed to analyse these data statistically, and so misses a number of points. First, none of the sections of the *Rep.* is statistically significantly different from any other (*Rep.* I = 16 ± 4, *Rep.* V-VIII = 20 ± 2, *Rep.* X = 23 ± 7, *Rep.* II-III = 24 ± 3) so there is no reason to enter them separately as he does. Second, no statistically significant grouping of dialogues can be made from the results. But third, the smooth trend is itself significant. Siebeck collected numerous types of assent and divided them into three (rather fuzzy) classes, and claimed to be able to see a weak trend away from problematic affirmatives (such as ὁμων) toward apodictic affirmatives (such as δῆλον). Siebeck acknowledged that other factors would influence this (p. 52) and B. notes that the classification itself is a problem (p. 53). Again B. canvasses various details, none of which can be supported from the data (p. 54).
Conrad Ritter, *Untersuchungen über Platon* (Stuttgart 1888), "was the first to write a book on the stylistic method" as applied to Platonic chronology, and B. rightly devotes over 30 pages to it (pp. 55-86). Ritter investigated the use of reply formulae and other expressions, which B. summarises in two dense tables (also included are Dittenberger's work on ὀπτερ etc. and Schanz on ὡντ- and ὡληθ-). B. notes the difficulty of counting such formulae (ἡληθη λέγεις appears also as α. μέντοι λ., τόδε μέν α. λ., πάνω μέν οὖν α. λ., etc.), and it is not always clear which variants to include; at any rate B. notes that Ritter apparently did not include all forms and counted those he did include erratically (pp. 55-6). Now the numbers are all relatively small (the largest single entry is 220 "Answers by means of a repetition" in Rep. and most are 5 to 10 times smaller), but neither Ritter nor B. seems to have thought of listing all passages cited (say in an appendix)—this would allow later workers to check and improve the results. As it is, all the work must be done anew; even B. only checked some results (p. 56). Ritter recognized the need to count occurrences with respect to the number of opportunities, but for reply formulae there is the difficulty that not every occurrence of some reply is an opportunity for any other reply (p. 57). So instead Ritter took as the base not the number of opportunities for a reply of similar meaning, rather all "reply formulae which are repeated with some frequency, together with a few which are sporadic but closely related to the former." Ritter sought to confirm the conclusions of Dittenberger and Schanz that Soph., Pol., Phil., Tim., Crit., and Laws form a late group, and came up with 45 results which seemed to show this. Some are clearer than others. For example (# 1) "ἐγὼγε, ἔμοιογε δοκεῖ"1 is rarer, while (# 2) "ὁρθὼς, ὁρθότατον (λέγεις), ἡληθετότατον" are commoner (p. 57). True enough but putting it so obscures the facts. From B.'s figures (Table 10.1, p. 58) we find that "ἐγώγε, etc." averages 16 ± 4 % of the total of reply formulae in Lach., Charm., Euthd., Crat., Euph., Gorg., and Meno (results for Prot., Apol., Crito, and Symp. show larger σ's and are hence excluded),2 while for Soph., Pol., Phil., Tim., Laws the average is 0.5 ± 0.4 %, effectively zero (Crit. figures are not given by B.). But for the Phdr. we find 1.4 ± 1.4 %, consistent with the late group (z = 0.6 or P = 46 % of difference), and very different from the early group (z = 3.4 or P' = 1 in a thousand); while for Phdo., Theaet., and Rep. we find 6.3 ± 1.1 %, very different from the 'early' group (P' = 2 %), and far more different from the late group (P' about 1 in a trillion); and the Phdr. is far less likely to belong to this 'middle' group (P' about 1 in a thousand) than to the 'late' group. Thus we find the proper grouping based on ἐγώγε, etc.:3

'Early' (Lach., Charm., Euthd., Crat., Euph., Gorg., Meno) 16 ± 4%

---

1 So in the Table 10.1; B.'s text p. 57 adds "ὁμων, etc."; which is it?
2 The σ's calculated as usual; for Prot. we have 6 ± 4 %, Apol. 10 ± 10 %, Crito 10 ± 7 %, Symp. 11 ± 6 %; the σ² on the overall average has been increased by a factor of 7 (the number of items averaged) to account for the combining of probably correlated data.
3 Prot. is more consistent with the 'middle' than with the 'early'; Apol., Crit., Symp. fall 'between' early and middle.
This is rather different than what one would gather from B.'s discussion (p. 57) and Table (10.1, p. 58). Similar remarks could be made for a number of other 'results'. For some of these one could readily figure the required ratios, for others it would be rather harder; for the remainder of the tests, the numbers are too small to yield much. Now comes the difficulty, faced but not solved already by Campbell, but far more acute for Riller—how to combine many different results which may not precisely agree? That is, what relative weight do we give to, say, the decreasing tendency to use ἔγωγε, etc., and, say, the shift from πάνυ γε to πάνυ μὲν οὖν? Ritter simply counts each feature as found or not found, and computes thereby a number which varies from 1 (Euph.) to 42 (Laws), except for Tim. and Crit. which are computed differently (pp. 66-7). In 1888, this may have been reasonable—a century later B. ought to have told us that such things could be properly done by any of a number of techniques of multivariate analysis (as in c. 22—see below). Ritter was also concerned to establish the unity of the Rep. (pp. 67-74, 79-81), and the chronological order of dialogues within his three groups (pp. 74-8, 82-3), and the question of the suspected works (pp. 83-6). Authenticity is for us here a separate question; while B. notes (p. 77) Ritter's figures are generally not adequate to make such fine distinctions. Ritter concluded that, so far as his number of characteristic expressions method could tell, the ten books of the Rep. are more or less identical (B. confirms this by taking pairwise averages to reduce variation due to varying book-length—effectively a moving average)—both Ritter and B. allow for the possibility that Book I may differ (i.e., be earlier).

J. Tiemann, "Zum Sprachgebrauch Platos," Wochschr kl Phil 6 (1889) 248-53, 362-6, 556-9, "Einige formalhafte Wendungen bei Plato" idem 586-9, and a review of Ritter, idem 791-7, 839-42, wrote primarily to correct and clarify Ritter (pp. 87-91). Most of his results are details depending on small numbers (as B. notes p. 89), though the use of superlatives in reply formulae seems pretty clearly higher in Phdo., Theaet., Rep., Soph., Pol., Phil., and Laws than in other works (so the Table 11.1, p. 88, though Tiemann and B. put it differently and not quite rightly).

G. B. Hussey, "On the Use of Certain Verbs of Saying in Plato," AJPhil 10 (1889) 437-444, examined certain forms of verbs which Plato used to refer to something already said (in an argument) in the dialogue (pp. 92-5). He noted an increasing percentage of the use of ἐπὶρηθ— and ἐκαθορίζοντα, though again the actual figures suggest an order rather different than he (or B.) saw. Percentages from B.'s table (p. 92), using all the forms (as Hussey intended and as seems reasonable from the consistency of the different numbers), are as follows:

1 Even this assumes all the results could be expressed in consistent numerical terms, i.e., ratios or some other form.
2 Though B. only says "too small", he means "so small that the ω's are too large," to determine the order.
Gorg. = 2.6 ± 1.9 Theaet. = 13 ± 3 Rep. = 14 ± 7 Symp. = 14 ± 7
Phdo. = 18 ± 7 Phdr. = 21 ± 8 Soph. = 23 ± 7 Laws = 25 ± 3
Phil. = 27 ± 6 Tim. = 46 ± 13 Pol. = 48 ± 9
where the first (Gorg.) and the last two (Tim. and Pol.) are the two groups one
could validly separate from the others (B.'s own figures would give for the last
seven the order Soph., Phdr., Theaet., Phil., Laws; and Tim., Pol.—only
roughly the same).

Hans von Arnim, *De Platonis Dialogis Quaestiones Chronologicae* (Vorlesungsverzeichnis der U. Rostock für das
Winter-Semester 1896), investigated reply formulae, apparently unaware of
Ritter's work (pp. 96-114). He divided reply formulae into five classes: 1) emphatic adverbs with particle, 2) expressing agreement and admitting truth, 3) ad-
verb with ellipsis of verb, 4) verbs expressing seeming or supposing, with ad-
verb, 5) rhetorical questions. But the results fluctuated in a way inconsistent
with the then-agreed order of groups of dialogues. One could make an important
point (which B. fails to note): even granting that the Laws was Plato's last
work, how are we sure that Plato's style, in so far as it is measurable by
stylometry, changed monotonically over time till he attained the style of the
Laws? Yet Arnim and B. here reject the apparent ups and downs (themselves
only apparent given a previously-assumed or established order) in favor of the
monotone hypothesis. Warning bells ought to have rung (as they should have,
albeit more faintly, on the comparison of almost any pair of previous results),
for Arnim and for B. But B. only notes that "as regards the dialogues of the first
period Arnim picks and chooses his evidence" (p. 108)—important but not
sufficient. For the percentage of rhetorical-question replies (pp. 102-3) and for
the percentage of ποίερα-uel sim. replies (p. 104) B. gives figures: as usual
without uncertainties; but his Tables 13.1 (pp. 110-1) and 13.2 (pp. 112-4) al-
low us to compute the very necessary σ's. We find for rhetorical-question
replies:

<table>
<thead>
<tr>
<th>Period</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early</td>
<td>5.2 ± 2.2%</td>
</tr>
<tr>
<td>Middle</td>
<td>13.0 ± 2.4%</td>
</tr>
<tr>
<td>Late</td>
<td>23.1 ± 3.5%</td>
</tr>
</tbody>
</table>

The dialogues grouped together are statistically indistinguishable; the pairs
'Early'/Middle' and 'Middle'/Late' have a 98% probability of being significantly
different. In the second case, for ποίερα-uel sim., we find similar results (but
fewer dialogues give figures at all). I have no particular reason to believe that
any of these differences necessarily corresponds to chronology.

C. Baron, "Contributions à la chronologie des dialogues de
Platon," *REG* 10 (1897) 264-78, examined anastrophe of περί (pp. 115-
22). Baron wisely sought to calculate ratios of περί to περί (equivalently, ratios
of ΠΕΡΙ after to ΠΕΡΙ before its object), and eliminate from consideration
phrases in which anastrophe was thought to be impossible or rare (περί πολλοῦ
[uel sim.] ποιεῖτοθατ and περί in a prepositional phrase qualifying a substantive

---

1The σ²'s are increased by a factor of the number of dialogues combined; the Apol.,
Prot., Symp., and Tim. have zero rhetorical-question replies, and are excluded.
or substantival article). B. corrects Baron's figures at many points and agrees with Baron that Plato used πέρι more with passing time, but "several works have been placed in a group to which they do not belong on a mechanical interpretation of their frequency figures" (pp. 116-7). But B. places the dialogues in order as determined from the corrected percentages (of πέρι/πέρι-total), deducting only π. π. π.—again without σ's. When we compute these we find that the apparent smooth increase (Table 14.2, p. 121) is so smooth and the σ's sufficiently large that even the Rep. and the Laws are not certainly distinguishable. One could make some use of the results (by, say, one-dimensional cluster analysis—a technique familiar to B. from his work on Wishart and Leach, see below), and could one combine these results with independent measures (reply formulae, e.g.) by means of multivariate analysis they could add to the evidence, but as they stand it is probably going too far to say as B. does "the works which exhibit the highest incidence of anastrophe are precisely those placed by earlier investigators in the middle and late chronological groups."

Wincenty Lutosławski, *The Origin and Growth of Plato's Logic* (London 1897), c. 3 (pp. 64-193), a most unusual man, investigated 500 (sic) stylistic features, on the assumption that all of Soph., Pol., Phil., Tim., Crit., and Laws were late, in order to establish incontrovertibly a framework of Platonic chronology (pp. 123-35). B. wisely does not review every feature. Lutosławski's approach is, as B. notes, flawed by his very peculiar mathematical treatment. Lutosławski computes units of stylistic affinity, where a characteristic of the later style is given a value for a particular dialogue depending on how often it is found therein (once evaluates as 1, a few times depending on length of dialogue evaluates as 2, a few more times as 3, more than once every two Didot-pages as 4). E.g., to borrow from von Arnim, when a rhetorical question as a response to a direct question occurs twice in Crito it has a value of 2, but it evaluates as 3 in Gorg. where it occurs 16 times (but had it occurred only 5 times in Gorg. it would still have had a value of 3). Then the dialogues closest in 'value' to the Laws (etc.) are assumed closest thereto in time. There are other rules, but B. has made it clear that the procedure is arbitrary, and hence unreliable. What B. does not do is evaluate adequately the features suggested by Lutosławski. B.'s evaluation (pp. 132-5) does indicate a number of important and valid criticisms—but he leaves us with 161 "acceptable" characteristics, without telling us anything positive about them (only that they are not to be ignored). It is Lutosławski who seems to have coined the term "stylometry" (p. 130), so one at least of his contributions has endured. I suspect many (B. would allow "less than a hundred") of Lutosławski's proposed characteristics are worth reinvestigating.

P. Natorp, "Untersuchungen üuber Plato's Phaedrus und Theaeetet," *Arch Gesch Philos* 12 (1899) 1-49, 159-86, and 13 (1900) 1-22, attempted to refute the rough consensus on Plato's chronology that had emerged via stylometry (because it conflicted with his view of Plato's

---

philosophical development; pp. 136-52). Natorp made use, as had earlier workers, of vocabulary figures—which (as B. notes, p. 152) depends heavily on subject. His calculations, like those of Campbell, indicated 'lexical affinity' to the Laws, Tim., and Crit. by computing the number of words a given dialogue had in common with that set of assumed late dialogues, and computing a ratio per page (Didot, see Table 16.2). He also computed similar 'affinities' for all dialogues with each of them (thus approaching the idea of numerical taxonomy—see below). I would not spend so much time on Natorp as B. has done, since (however one proceeds numerically) there is a fundamental flaw in the procedure—vocabulary (so difficult to count in any case) is very context-dependent unless the stylometer chooses carefully indeed. Natorp did not publish the 1,949 words on which he based his work (p. 152). There is a better way, related to the problem of the unseen species in biological trapping problems, which has been applied to Shakespeare.¹

Walter Janell, "Quaestiones Platonicae," Jbb f cl Phil S. 26 (1901) 263-336, returned to hiatus in greater detail (pp. 153-66). His first task was to differentiate between permissible and impermissible hiatus, which he did primarily in a somewhat intuitive fashion (e.g., αὐτη η is permissible since their separation "would be somewhat harsh") and in a definitely circular fashion—those types of hiatus found considerably more often than the rest in Soph., Pol., Tim., Crit., and Laws are permissible (but is not that what Janell was trying to measure: Plato's changing avoidance of hiatus?). What he found, as B. makes clear, was that Plato considered hiatus within a phrase or at a pause objectionable, but not with καί, the article, περί, μη, δη, η with following, τι or τι, ἄν with preceding relative, ει with following, ὤ with following, πρό, and ει in combination with the following verb. In fact Janell and B. count instances of objectionable hiatus and note that the number per page ranges from 46 (Laches) smoothly down to 28 (Menex.), but is much lower for Laws, Phil., Epin., Tim., Crit., Soph., and Pol. (8 or 9 down to less than 1). Since the distinction between objectionable (within a phrase or at a pause) and allowable (with the words listed above) seems clear (whether or not Plato would have made it this way), and although the method of counting is flawed, the figures are so different that they are likely to be significant (but of what?). As usual B. does not evaluate them statistically. B. is aware that one ought to count with respect to possibilities for hiatus (pp. 162-5), and counts instances of the words listed above (καί, etc.) causing hiatus, and instances not doing so in Crat. and in Pol. (p. 164 = Table 17.7), chosen as they are "roughly the same length" (42.3 versus 43.2 Didot pages).² The figures seem to show a much stronger avoidance of even "allowable" hiatus; e.g., Crat. has a ratio of καί with hiatus to all καί of 0.41 ± 0.03, while Pol. offers 0.24 ± 0.02 (the article, περί, μὴ, δη, η, τι, τι,

²Why not Soph. 39.6 versus Prot. 39.5, or Symp. 39.3?
öτι (conj.), ὅ, and verbs in -θατ give very similar figures for Crat.; Pol. is much lower for most, as high as 0.19 ± 0.04 only for περι). As B. notes (p. 162) this might very well alter the results, and so he tabulates (for Hipp. Ma., Ion, Menex, Phdr., Laws, Phil., Epin., Tim., Soph., Crit., Pol. only) the incidence of "permissible" hiatus (and elision which "may have originated in transmission"): Table 17.6. He concludes this does not alter the order (rather it does not alter the grouping), since here apparently "permissible" hiatus does not vary greatly.\(^\text{1}\) It is too bad—one would think hiatus would be the most help, and all we are left with is a muddle.

B. departs from his chronological order to consider two investigators of clausulae who arrived at similar results: W. Kaluscha, "Zur Chronologie der platonischen Dialoge," WSt 26 (1904) 190-204, and L. Billig, "Clausulae and Platonic Chronology," JPhilol 35 (1920) 225-56 (pp. 167-206). Only here does B. show knowledge of correct statistical procedure—and he has published (with a statistician) a paper on this very subject: D. R. Cox and B., "On a Discriminatory Problem Connected with the Works of Plato," J Roy Stat Soc (B) 21 (1959) 195-200. An advantage of counting clausulae is that each sentence is a possible instance of each clausula, so that the adjustments necessary even for hiatus (above) and crucial for vocabulary (above) are absent here. This allows more straightforward comparison of counts. Kaluscha examined clausulae of 5 syllables which makes 32 different varieties (and he excluded a few passages whose scansion might be ambiguous); but Kaluscha concluded that in the Laws the final syllable (normally anceps) could be considered always long, which reduced the number to 16 varieties there. He claims that in the Laws Plato favored five clausulae:\(^\text{2}\) II.4 = - - - - - , III.9 = - - - - - , III.4 = - - - - - , II.10 = - - - - - , V = - - - - - . These, he says, make up roughly 55% of each book of the Laws, and so are far more frequent than the other 11 possibilities. There are other details (always the five most common clausulae are noted) and a mass of tables, but B. scouts the possibility of any certain conclusions (p. 180) due to uncertainties in Kaluscha's procedure. In fact whether or no the procedure and counts are accurate, the statistics do not allow much in the way of conclusions.\(^\text{3}\) When, e.g., the 32 clausulae of Soph. and Tim. are plotted as a histogram, it turns out that there is no distinction statistically significant, save that in Tim. III.3 = - - - - - and I.2 = u-u-u are slightly preferred. If however the clausulae are considered as having final anceps

\(^\text{1}\) In fact the presentation of the data is confusing — if "Class I" is permissible (Table 17.1 and p. 155), how can the Pol. have 308 such in Table 17.1, but 694 "permissible" in Table 17.6? (And so likewise Soph., Crit., Tim., Phil., and Laws, the other five dialogues appearing in both tables.).

\(^\text{2}\) I follow B.'s notation temporarily for simplicity of account, but it is ambiguous and arbitrary; clearer would be to let - = 1, u = 0, and label each clausula by its binary-numeral; thus "II.1" = -u-u becomes 11000 or 24 converting to ordinary decimal numerals. Cox and B. suggest the first part of this method, p. 195.

\(^\text{3}\) For Crito and Crit. the numbers are too small; they are marginal for the Apol.
and the resultant 16 clausulae plotted as a histogram, then in Soph. (III.1 = - - - v-v) + (III.4 = - - - v-v) is slightly preferred, while (II.6 = v- - v-v) + (III.8 = - - - v-v) is slightly avoided; but in Tim. (III.2 = - - v-v) + (III.3 = - - v-v) is slightly preferred. In Phil. and Pol., at first glance showing more spread between clausulae, the results are similar: Phil. prefers III.4 = - - - v-v and Pol. prefers V = - - - - -; considering the final syllable as anecps, we find that Phil. prefers (III.1 = - - - v-v) + (III.4 = - - - v-v) and (II.5 = v- - v-v) + (III.9 = v-v - v-v) and Pol. prefers (III.1 = - - - v-v) + (III.4 = - - - v-v). There is certainly nothing special about the "top five" clausulae—it is merely an arbitrary selection, and rightly one ought only to select those (top or bottom) which are statistically significant. When we examine the remaining dialogues which are long enough to give good statistics (still using Kaluscha's uncorrected numbers), it is very difficult to find a simple linear chronology. So all of B.'s elaborate lucubrations, in some cases statistically quite elegant, are I believe rendered moot by the data. The conclusion (p. 206) that "it seems unreasonable to doubt that the sequence [is] Tim., Crit., Phil., Pol., and Laws" is untenable, and the fundamentally naive method of noting in each dialogue the five most frequent clausulae (of 32 or of 16; compare pp. 168-79 following Kaluscha) is insupportable. Cox and B.'s attempt (B. pp. 198-206) to model the changes between the Rep. and the Laws in usage of all 32 clausulae is statistically elegant but flawed because it assumes the change is monotone (Cox and B. pp. 195, 6), but the data for the dialogues placed between (Crit., Phil., Pol., Soph., Tim.) show that it was not so (Cox and B., p. 198—B. Table 18.11a, p. 205). It is not just that the order Cox and B. favor does not fit, in fact no possible order fits that assumption. Again a sore disappointment—one would think that clausulae ought to give as clear an indication as anything, yet despite untangling this mare's-nest of wrong assumptions and poor statistical procedure, we are left with more results than answers.

H. von Arnim, "Sprachliche Forschungen zur Chronologie der platonischen Dialoge," Sitzber der kais Akad der Wiss in Wien, Phil-Hist Kl 169.1 (1912), aimed to endow the stylometric conclusions about Platonic chronology with such "conclusive force that any opposition would be impossible" (pp. 207-220). Even in philology there is no immovable object, whether or not it is weighted down with elaborate algebra, as in this case. Von Arnim sought to compare the use of pure reply formulae (those not repeating any word from the question) in every pair of works (counting books of the Rep. and Laws as separate and excluding Apol., Menex., Tim., Crit., and Laws V, XI due to lack of dialogue). In order to make this comparison (which properly ought to be done by some version of multivariate analysis, see below), von Arnim invented an arbitrary affinity value formula, which B. discusses at length. The results are (after overmuch algebra): Ion, Prot. ... Lach., Rep. I, Lysis, Charm., Euph., Euthd., Gorg., Meno, Hipp.Mi., Crat. ... Symp., Hipp.Ma., Phdo., Crit. ... Rep. II-X, Theaet., Parm., Phdr., Soph., Pol., Phil., Laws (where ellipsis indicates lapse of time). But the magic formula is all an illusion (as B. shows, pp. 216-7) even on von Arnim's assump-
tions, and in any case is arbitrary and unfounded. I should not have spent so much time as B. does in revealing that there is nothing behind the curtain.

C. Ritter, "Unterabteilungen innerhalb der zeitlich ersten Gruppe platonischer Schriften," Hermes 70 (1935) 1-30, investigated the use of μήν, ὥς with superlative, ἀλλος/ἐτερος, ὅςος, ὁσπερ, and ὁν, and very rightly insisted on distinguishing uses by sense (pp. 221-7). But the evidence seems frail, as most of these are used but rarely in the early works (Table 20.1 lists 67 total uses of all items considered in twelve dialogues, Rep. I being counted as one). B. wisely concludes, "considering how low the figures are," that here the argumentum ex silentio is dangerous, and the absence of one or another of these uses from a dialogue is meaningless.

A. Diaz Tejera, "Ensayo de un metodo lingüistico para cronología di Platón," Emerita 29 (1961) 241-86, studied non-Attic vocabulary, as documented in the Koiné, in its incidence in Plato (pp. 228-34). He counts the number of examples of various categories on non-Attic elements (neologisms, ionicisms, ionic and formerly poetic words, etc.) per page, and claims that larger values indicate later works. As B. amply shows, his conclusions are often not supported by his own data, his classification of words is often arbitrary, and even his figures are unreliable. B. ought to have added that the assumption that the Greek dialects develop uniformly into Koiné is not necessarily so.

David Wishart and Stephen V. Leach, "A Multivariate Analysis of Platonic Prose Rhythm," Computer Studies in the Humanities and Verbal Behaviour 3 (1970) 90-99, two statisticians, attempted to characterise not just five-syllable clausulae, but whole sentences (in overlapping five-syllable groups: "1-5, 2-6, 3-7, and so on," Wishart-Leach p. 90), in 33 samples from ten works: Tim. (9 samples), Soph. (1), Phil. (1), Crit. (3), Laws (2 from VIII, 3 from V), Epistle VII (1), Rep. (1 from II, 2 from X, the myth of Er), Pol. (1), Phdr. (5) and Symp. (4) (pp. 235-48). Their method, which B. explains rather well (thanks to the statistician F. J. W. Rendell—see p. 246, n. 12), is essentially to compute for each pair of samples the sum of the squares of the differences in corresponding percentages of the 32 five-syllable rhythms. Then these distances are clustered by successively adding those samples to a group whose distance from the group average is minimum.¹ This clustering does not always reproduce the expected grouping: one Phdr. sample was clustered with the Symp. samples and the two Rep. X samples, while the Tim. and Crit. cluster included the Soph. sample as well as the Rep. II sample. Since the Phdr. sample grouped with the Symp. was Lysias' speech, the statisticians "conclude tentatively that it was a genuine speech of Lysias" (p. 241). Whether it is or not I do not know; what I do know is that the conclusion does

not follow from the evidence. Moreover *Tim.*/*Soph.* clustered with *Crit.* at the next level of significance, while *Phdr.* clustered with *Syrp.* + *Rep.* X (+ "Lysias"); and at the third level *Laws* (with which *Phil.* clustered) was closer to *Tim.*, *Crit.*, *Soph.* (+ *Rep.* II). Such a cluster analysis suggests similarities not chronology, so Wishart and Leach turned to two other techniques of multivariate analysis, "principle components" and "multidimensional scaling". Essentially each of these tries to map or project many variables onto a few in order to clarify the data (in a similar way, the surface of the Earth is mapped, by geometric projection, onto a flat sheet of paper). They proceed by finding that lower-dimensional space through the original space (here the 32-dimensional space of five-syllable rhythms) which when the points are projected onto it minimises the root-mean-square error. The hope is that there will exist some mapping which does not too much distort the data but renders it comprehensible. In principle components analysis, an iterative algebraic process identifies some combination of the original variables that 'explains' the observed variation, just as, say, sociological data is often found to depend mainly on a few factors such as income, gender, etc.; multidimensional scaling is more like the geometrical projection just described. Their similarity is neatly revealed by the very similar results in Figures 4 and 5 (pp. 244-5, from Wishart and Leach): the first displays the data in terms of the first two principle components, the second as a 2-dimensional projection by multi-dimensional scaling. Again *Rep.* and *Syrp.* are close, *Tim.*, *Crit.*, *Soph.*, and *Pol.* are close, and *Phil.* and *Laws* are close. Hardly surprising since the data being analysed in each case are the same set as gave the distances used in cluster analysis above. But the results do not allow an unambiguous linear (i.e., chronological) ordering—the variation is clearly at least two-dimensional, and it too much strains the data to produce the one-dimensional plot (p. 246), though it be mathematically possible. B. is quite right to criticise Wishart and Leach's conclusions regarding the *Phdr.*—"all five samples of this work were taken ... from speeches" (p. 247), which biased the data.

Summary and Prospect

---


2 Visually, we could imagine the leaves on a tree as 3-dimensional points, and the fitting or projecting as the finding of a plane through the leaves onto which they could be projected (as a map) with the minimum 'distortion' (minimum rms); for a birch, this might be a plane through the trunk, for a banyan, a plane perpendicular to the trunk.

3 In fact the multi-dimensional scaling technique generates a measure of strain (called "stress"). In the two-dimensional picture it is 0.038 (rather small); in the one-dimensional picture it is 0.384 — quite large, and not consistent with a one-dimensional model (i.e., not consistent with assuming the one-dimensional picture is at all an accurate representation of chronology, or of anything).
We may conclude this analysis of B.'s book by noting that (1) he provides a very useful (if slightly flawed) survey of past results, which is strong on summary, (2) his analyses are often on target (especially when the work disagrees with the consensus) and are necessary reading for any future stylometrician, but (3) he fails to analyse in a sufficiently rigorous statistical way the results which these earlier workers do provide, a failure which becomes acute with Janell's study of hiatus and Kaluschka's study of clausulae, and this leads him to be more optimistic than is warranted, and (4) he misses completely (as did all the workers he surveyed) an important methodological point to which I shall come below.

The news seems bad—a survey of all the work done to date more or less proves that no one knew what he was doing, and even the censustaker has got his statistics muddled, while the latest book ( Ledger)\textsuperscript{1} using the full machinery of the computer age produces what will not stand up to inspection. What have we learned since Schleiermacher? From a battlefield strewn with stylometries dead and dying I pick out some important methodological lessons.

First, two equal and opposite mistakes. The stylometrician must be aware of what Pythagoras said and of what Aristotle said. Some (I think of Ledger or poor von Arnim [1912]) have become so caught up in the machinery that they forget the limitations of the material. Oppositely, what mathematics and statistics are used ought to be appropriate—the repeated failure to use even the simplest tests of significance (excusable in 19th-century philologists) is not acceptable.

Secondly, the difference between significance and meaning. A measure (say of hiatus) may be statistically different with some stated numerical significance (e.g., "at the z = 5.9 level" or equivalently "with a P' of about 4 in a billion") and still not be meaningful. Significance is a statistical (i.e., numerical) concept, while meaning refers to relevance. Brandwood seems to have dimly perceived this (p. 86), Ledger a bit more clearly (p. 4: the variables measured must have "a fair chance" of being "linked to stylistic features and not be just measurements of random and haphazard events"), but he does not seem to follow through (despite coming close, see pp. 48, 51).

Thirdly, two mathematical assumptions not much appreciated. It is not always (is it ever?) true that any measurable linguistic feature will display monotone changes with time or uniform changes between authors. With respect to chronology, both Brandwood (p. 78) and Ledger, again more clearly (pp. 171-6), though again he falls short in the throw, have recognized this. It must be stated clearly and kept always in mind (especially with chronology)—the relationship between the answer sought and the feature measured is likely not linear. Moreover, great care must be exercised when attempting to combine different measures, whether or not they are like in result. When this is done (to return to the Aristotelian pole) it must be done in such a way that the result is not only statistically correct (and hopefully significant) but philologically meaningful. That is, it must be possible to reify (i.e., in some way to express what we are being told by) the mathematics. When archaeologists classify, say, pots by this

\textsuperscript{1}Reviewed in \textit{BMCR} 2 (1991) 422-7.
method, they can in the end describe what the mathematics has told them about
the differences between their pots.

Fourthly, to deepen Aristotle's point (if I may make so bold)—neither
should we be a priori complete sceptics about stylometry, nor should we expect
a priori to answer by stylometry all questions with mathematical precision.
Stylometry studies people not physical particles and the mind is self-aware,
neither random nor determined. Authors vary, even significantly, and we must
expect that. Minds change, both in knowing and in writing, and necessarily so,
and therefore by a kind of Heisenbergian uncertainty principle of epistemology
we as minds can never know with ultimate precision any other mind, especially
one long dead who wrote in Greek, and so brilliantly that his political theories
still seduce millions. Yet at the same time, we may, like Sherlock Holmes, be
able by careful examination of the evidence to say a thing or two (even Moriarty,
though genius, was susceptible of detection). Just because neither random nor
determined, the mind is patterned or has character, which can be partially
known—and perhaps measured?1 Although Gulliver encountered sages in
Lagoda who sought new wisdom in random letters and although W. A. Mozart
left careful instructions on how to compose 'as many German waltzes as one
pleases' by casting dice (Koechel 294D),2 most authors display certain regulari-
ties.

There is a fifth methodological lesson and one I would argue is in a practical
sense the most important. Aristotle advocated proceeding from the known to the
unknown (e.g. Phys. 1.1 [184a17-25]) as a necessity for the advancement of
knowledge. All good modern science, whether physical or philological, does
proceed in this way. Yet no stylometrician that I know of, save one, has
proceeded in this way. All attack the unknown directly, and so seemingly hope
to reverse Aristotle in his own court. This is the worst muddle of all. But let
me now praise a neglected book—wrongly neglected, for it is the one example of
the correct procedure.

Milton Perry Brown, *The Authentic Writings of Ignatius* (Duke

Although well-received by reviewers, it has since languished. It should be re-
quired reading. Brown very wisely did not attack the unknown by the unknown,
rather he took a corpus in which there is now (since the work of Bishop Ussher
and Isaac Voss in the XVII cent. and of Bishop Lightfoot in the XVIII cent.)
 scholar unanimity that one part is authentic, the other by a much later writer.
But the later writer was imitating Ignatius' style—can we, knowing which is
which, find stylometric tests which make the distinction correctly? This is the
question Brown set himself to answer (with a view to applying the results to the

---

1 For similar points cp. Thos. Fleming, "Science and Philology," unpublished essay,
2; and the remarks of I. J. Good in the discussion (pp.225-7) to the paper by
233 (A reference I owe to Wm. Kerr), @227.

Pauline corpus)—essentially he has performed a necessary control experiment. He tries various tests (vocabulary, prepositions, particles, clauses, use of moods, tenses, and voices) and notes which ones work and which ones do not (sentence-length and word-order, for example, do not, while Ignatius' fondness for certain compounds does). Brown alone since Schleiermacher has seen that one must test the tests ("run blanks") in order to determine which are meaningful. In the absence of such blank runs, no sorites of results, however significant, can claim meaning.

And the chronology of Plato?—I know no more than Campbell; nor does anyone. A Campbell would give the Scottish Verdict: "Not Proven."¹

Paul Keyser
University of Alberta

* * * * * * *
I am indebted to Hugh G. Robinson for first introducing me to this fascinating topic, many years ago (1978); to Wm. M. Calder III for encouraging me in it over the years and for suggesting this review; to Richard Hamilton for giving me the great opportunity to write it; and to E. C. Kopff and E. E. Schütrumpf for stimulating discussions. But none of these is responsible for my scepticism or method.