ASPECTS OF LATIN

Papers from the
Seventh International Colloquium
on Latin Linguistics

Jerusalem, April 1993

Edited by
Hannah Rosén

SONDERDRUCK

Innsbruck 1996
How to do things with words per strong stop
Two studies on the Historia Augusta and Cicero

BERNARD FRISCHER

Deux études sur l'histoire Auguste et Cicéron: En étudiant les variations de longueur des «phrase» (c.-à-d., le nombre des mots par chaque signe fort de ponctuation), Marriott a tenté de prouver qu'un auteur seul a écrit l'Histoire Auguste. Notre étude soulève des objections philologiques et montre que l'analyse de Marriott comprend aussi des imperfections statistiques. L'étude d'autres historiens et biographes nous révèle quelque chose d'inattendu: le nombre de mots par signe fort de ponctuation n'est pas un bon vérificateur d'auteur, puisqu'il est assez constant d'un auteur latin à l'autre. Cette homogénéité est rattachée à la théorie de rhétorique au sujet de la période. La théorie ciceronienne nous suggère que la variabilité de longueur de la période de l'orateur sera plus grande que celle de l'historien. L'étude des discours de Cicéron soutient cette conclusion. La théorie ciceronienne nous suggère aussi que les discours écrits dans le genus grande devraient avoir plus de mots par signe fort que ceux écrits dans le genus humile. Les deux discours cités par Cicéron lui-même (Pro Caec. et Pro Rab.) en sont une bonne démonstration.

1. “Sentence”-length and the Historia Augusta: A philological and statistical critique of Marriott and of Sansone's critique of Marriott

The so-called Historia Augusta is a late-antique collection of biographies of Roman emperors, heirs-apparent, and pretenders from Hadrian to Numerian. For the long period it covers (A.D. 117–285), it often is our main source of information about imperial history. Its singular importance is therefore clear: as Sir Ronald Syme wrote, “the student of the Roman Empire cannot do without the HA” (Syme 1983a: 12). According to the manuscripts, the various lives are the work of six different authors. Dedications of thirteen of the lives to Diocletian and Constantine suggest a date of composition in the period ca. 290 to 325.

Ever since publication of a celebrated article by Hermann Dessau in 1889, these “facts” about authorship and date have been a matter of lively dispute. Arguing primarily on the basis of nomenclature and style, Dessau proposed two theses: (1) that the collection was not written by six different authors but by one only; and (2) that this author was active, not in the period of Diocletian and Constantine, but much later in the fourth century. Proving, disproving, or amending one or both of Dessau’s theses has been the focus of Historia Augusta research ever since.

In this effort, one work stands out as offering seemingly objective evidence in favor of Dessau’s first thesis: Marriott 1979 presents two studies that used

1 Dessau 1889 and 1892.
stylometrical arguments to discriminate the author of the *Historia Augusta* from other authors datable to the fourth century. The result appeared to confirm the Dessau-thesis of single authorship, and this was greeted with enthusiasm by several scholars, including Sir Ronald Syme, who wrote: “it will not be easy to impugn [this study]. Those who conducted the inquiry may await attempts without undue apprehension.”

My first goal is to show that it is indeed rather easy to impugn Marriott’s first (and more important) study, on both philological and statistical grounds. Then I will consider ancient rhetorical theory and oratorical practice regarding proper “sentence”-length. The two goals are intimately related, for Marriott’s first study was precisely an attempt to distinguish the *Historia Augusta* from other fourth-century texts on the basis of “sentence”-length. Now, “sentence”-length is a concept that depends on how one punctuates a text, and different editors will do this differently. A compound sentence whose parts are separated by a comma or semi-colon in one edition may be divided into two independent sentences in another edition. Thus, in quantitative studies it has long been recognized that a higher degree of accuracy will be achieved if we count, not words per sentence, but words per strong punctuation mark, which is basically any of the above except the comma.

Marriott defines “sentence”-length in the *Historia Augusta* as “a sequence of words terminated by a full-stop, colon, or interrogation mark.” He compared the distribution of word-sequences in the *HA* to those found in the following control texts: the anonymous *De rebus bellicis*, the Codex Theodosianus, and Ammianus Marcellinus. As can be seen on table 1, the result of the comparison seemed to be a philologically and statistically significant difference between the distributions of these texts.

<table>
<thead>
<tr>
<th>TEXT</th>
<th>MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Historia Augusta</em></td>
<td>16.06</td>
</tr>
<tr>
<td><em>De rebus bellicis</em></td>
<td>21.73</td>
</tr>
<tr>
<td><em>Codex Theodosianus</em></td>
<td>26.19</td>
</tr>
<tr>
<td><em>Ammianus Marcellinus</em></td>
<td>36.00</td>
</tr>
</tbody>
</table>

Table 1: Mean “Sentence”-length in the *HA* and some fourth-century texts

[Source: Marriott 1979]

I emphasize the word “seemed” in the previous sentence, for Sansone 1990 pointed out some of the fundamental problems with Marriott’s study. First, Sansone rightly

---

2 Syme 1983b: 212. Marriott’s conclusion has also been accepted by Honoré 1987: 156 n. 8 and Scheithauer 1987: 160.
4 Marriott 1979: 66. In what follows, I put the word *sentence* into quotation marks when the word is used loosely to mean “a sequence of words between two strong stops.”
5 Sansone’s critique has been accepted by Meißner 1992.
criticized the failure to include hyphens and semi-colons among the strong stops indicating what we can loosely call "sentence"-end. We might also add that words in parentheses should also have been taken into account. Secondly, Sansone noted that Marriott's control texts represent different literary genres: the *HA* is biographical; the *De rebus bellicis* is a technical treatise; the *Codex* is a legal code; and Ammianus wrote a history. Sansone observed that the only control text Marriott used from the same genre as the *HA* was the Liber *De Caesaribus* of Aurelius Victor, but Marriott inexplicably used Victor only in his second study of word-type at the beginning and end of "sentences", not in his comparisons of "sentence"-length. Since generic conventions affect so much of ancient style, Sansone rightly considered this comparison of apples and oranges to be a major flaw in the design of Marriott's "sentence"-length experiment.

Sansone's purpose was purely critical; he did not attempt to repeat Marriott's "sentence"-length study, making the necessary corrections. This we can do by comparing the *HA* to other examples of Roman biography and historiography, including this time the text of Aurelius Victor. As table 2 shows, with proper control texts, the large differences Marriott found between the *HA* and other texts all but vanish.

<table>
<thead>
<tr>
<th>AUTHOR/TEXT</th>
<th>CENTURY</th>
<th>GENRE</th>
<th>MEAN WORDS/STOP</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historia Augusta</td>
<td>IV</td>
<td>Biography</td>
<td>16.06</td>
<td>10.67</td>
</tr>
<tr>
<td>Eutropius</td>
<td>IV</td>
<td>History</td>
<td>14.42</td>
<td>07.77</td>
</tr>
<tr>
<td>Aurelius Victor</td>
<td>IV</td>
<td>Biography</td>
<td>16.36</td>
<td>09.19</td>
</tr>
<tr>
<td>Suetonius</td>
<td>II</td>
<td>Biography</td>
<td>17.18</td>
<td>10.32</td>
</tr>
<tr>
<td>Tacitus, Ann., Hist.</td>
<td>II</td>
<td>History</td>
<td>14.06</td>
<td>08.79</td>
</tr>
<tr>
<td>Tacitus, Agricola</td>
<td>I</td>
<td>Biography</td>
<td>14.93</td>
<td>09.69</td>
</tr>
<tr>
<td>Livy</td>
<td>I B.C.</td>
<td>History</td>
<td>14.92</td>
<td>10.35</td>
</tr>
<tr>
<td>Caesar</td>
<td>I B.C.</td>
<td>History</td>
<td>17.33</td>
<td>11.87</td>
</tr>
<tr>
<td>Cornelius Nepos</td>
<td>I B.C.</td>
<td>Biography</td>
<td>15.87</td>
<td>10.17</td>
</tr>
<tr>
<td>Sallust</td>
<td>I B.C.</td>
<td>History</td>
<td>15.90</td>
<td>10.26</td>
</tr>
</tbody>
</table>

Table 2: Mean words per Strong Stop in *HA* and generically-related control texts

From these data, it is clear that the mean length in the *HA* (16.06) is near the middle of what may be called the normal range for biography and history. This extends from a low of 14.06 (Tacitus' historical works) to a high of 17.33 (Caesar). Mean length

---

6 SD = standard deviation. The texts used were the following: Eutropius — random sample of 100 sentences (ed. Ruehl [1919]); Aurelius Victor — random sample of 100 sentences (ed. Pichlmayr [1911]); *Historia Augusta* — complete text (Hohl 1965); Suetonius — complete text (Ihm 1908); Tacitus, Agricola, Ann. 1–3, Hist. 1–3 (Fisher 1906); Livy — random samples of Books 1–4, 24–26 (Conway, Walters 1955); Cornelius Nepos — complete lives of Alcibiades, Atticus, Chabrias, Cimon, Conon, Dion, Lysander, Miltiades, Pausanias, Themistocles, Thrasybulus, Timotheus (Marshall 1977); Sallust — complete *Bellum Jugurtinum* and *Caietinae* (Kurfess 1957).
does not appear to be determined by the difference in genre between history and biography, nor by the difference in date. The range among the historians is from 14.06 (Tacitus) to 17.33 (Caesar). Among the biographers, it is quite similar, going from a low of 14.93 (Tacitus' Agricola) to a high of 17.18 (Suetonius). In the first century B.C., the means vary from a low of 14.92 (Livy) to a high of 17.33 (Caesar); in the first and second century A.D., the variation is similar, rising from a low of 14.06 (Tacitus' historical works) to a high of 17.18 (Suetonius). Finally, in the fourth century, we find much the same thing: the lowest mean is 14.42 (Eutropius); the highest is 16.36 (Aurelius Victor).

We should also note that the difference Marriott found between the HA and one of his control texts — Ammianus Marcellinus — was inflated by the fact that Marriott used C. U. Clark's edition (Berlin 1910) that consciously applied a peculiar system of punctuation favoring the comma over the other internal punctuation marks. Using Gardthausen's edition of 1874 or Galletier's of 1968 (both of which use a more standard punctuation) makes the mean for Ammianus drop from Clark's 36.0 to 22.7 and 23.4 respectively. (Here it might be noted with regret that Marriott never specified which editions he used nor why he chose the ones he did.) Now the adjusted difference between Ammianus and the other texts is still stylistically significant. In explaining this difference, we should perhaps take into account the fact that Ammianus was a native speaker of Greek, not Latin, and his Sprachgefühl may therefore be abnormal. Because, e.g., Greek has the definite article and uses particles with greater frequency, Greek sentences tend to be longer than Latin ones. For example, in Book I of Polybius the mean is 25.0 (S.D. = 13.65; random sample of 250 sentences; text: Paton); in Plutarch's life of Cicero the mean is 24.6 (S.D. = 14.73; text: Ziegler); in Book 50 of Dio the mean is 24.5 (S.D. = 16.67; text: Boissevain).

Table 3 shows that within the HA itself, we find practically the same distribution of mean lengths as we find between the HA and the control texts used in table 2:

<table>
<thead>
<tr>
<th>Author/Text</th>
<th>Century</th>
<th>Genre</th>
<th>Words/Stop</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historia Augusta (whole)</td>
<td>IV</td>
<td>Biography</td>
<td>16.06</td>
<td>10.67</td>
</tr>
<tr>
<td>Spartianus</td>
<td>&quot;</td>
<td>&quot;</td>
<td>16.03</td>
<td>09.90</td>
</tr>
<tr>
<td>Capito linus</td>
<td>&quot;</td>
<td>&quot;</td>
<td>16.54</td>
<td>10.60</td>
</tr>
<tr>
<td>Gallicanus</td>
<td>&quot;</td>
<td>&quot;</td>
<td>14.24</td>
<td>09.64</td>
</tr>
<tr>
<td>Lampridius</td>
<td>&quot;</td>
<td>&quot;</td>
<td>15.99</td>
<td>11.07</td>
</tr>
<tr>
<td>Pollio</td>
<td>&quot;</td>
<td>&quot;</td>
<td>16.61</td>
<td>11.65</td>
</tr>
<tr>
<td>Vopiscus</td>
<td>&quot;</td>
<td>&quot;</td>
<td>15.37</td>
<td>10.40</td>
</tr>
</tbody>
</table>

Table 3: Mean words per Strong Stop in HA (whole) and in the six HA authors

Here, the lowest mean is found in Gallicanus (14.24); the highest occurs in Pollio (16.61). Thus, if Marriott's assumption that words per strong stop is a good indicator of authorship were correct, we could at this point use the data of tables 2 and 3 to reject Dessau's thesis of unitary authorship of the HA. The reason is clear: if Marriott is
on the right track, Dessau’s thesis would require that the range of lengths within the \textit{HA}\textsuperscript{-}group be much smaller than that in the entire group of control texts; but this is not the case. In fact, the \textit{HA} range of 14.24 to 16.61 is quite similar to what we found on table 2. The Dessau-Marriott thesis would also require that the overall \textit{HA} mean (16.06) differ markedly from the overall mean of the control group (15.49 excluding the \textit{HA}, or 15.70 including it). However, the difference we actually find of about one-half word is philologically trivial – and far less than the difference that emerged when Marriott’s generically unrelated control texts were used. Here it is important to distinguish sharply between statistical and philological significance. Given a large enough sample, the difference between a mean of 15.49 and 16.06 could easily be statistically significant in a hypothetical case. This would simply indicate that the difference is not the result of chance. It does not imply that the difference has any particular importance from the point of view of philology or stylistics. To gauge such importance, there is no substitute for the judgment of the qualified subject-expert. Thus, even when we utilize statistics, a subjective element of judgment remains inevitable and, indeed, necessary in our work as philologists.

At this point, we should consider Sansone’s third criticism of Marriott. Implicit in Marriott’s study and in my preceding remarks is the assumption that number of words per “sentence” is a legitimate tool of analysis. Sansone challenges this assumption head-on. He grants that it might be valid for works of modern literature, which were punctuated by their authors. However, he denies that it is valid for ancient Latin texts punctuated by modern editors because there must inevitably be differences between the modern text and the author’s lost original, just as there are (allegedly) great differences in punctuation between modern editions of an ancient text.\footnote{Cf. Sansone, 1990: 174–75.}

On this score, I believe that Sansone has gone too far in his critique, exaggerating the difficulties that arise because we must rely on modern editions. I would defend Marriott’s reliance on words-per-strong-stop on five grounds. First, as Wingo has shown, Latin literary texts were typically punctuated. The units indicated by punctuation marks include: word-breaks that we, today, mark with spaces; phrases and clauses that we, today, mark with commas; compound sentences that we mark with semicolons; parenthetical asides that we mark with hyphens or parentheses; sentences that we mark with periods, exclamation points or question marks; and paragraphs, which we, today, indicate by indentation.\footnote{See Wingo 1972: 94–131.} Thus, in studying words per strong-stop, we are working with a feature that falls within ancient practices of punctuation.

Secondly, as Wingo has also shown in the cases of texts like the \textit{Monumentum Ancyranum} and the \textit{Laudatio Murdiae}, when an ancient literary text happens to survive, there is a high degree of coincidence between ancient and modern punctuation.\footnote{Wingo 1972: 37–41 (Monumentum Ancyranum); 90–91 (Laudatio Murdiae).}
Thirdly, Sansone has overestimated the degree of difference to be found in two modern editions of the same text. I have compared mean words per strong stop in two editions of Ammianus Marcellinus produced a century apart by scholars of different nationalities. I found that there was only a slight difference in the two means. The study considered random samples of about 200 "sentences" of Books 14–16 of Ammianus Marcellinus. As noted earlier, in Gardthausen's edition of 1874, mean number of words per strong stop was 22.7. In Galletier's text of 1968, the mean was quite close: 23.4. Similar results are also found in a study of two editions of Cicero's Sixth Philippic. Fourthly, this similarity undoubtedly arises in part from the fact that in Latin there are some obligatory lexical markers of sentence-beginning, especially the inceptive adverbs.

Finally, Sansone does not take into account the fact that statistical tests typically have confidence intervals and so are robust enough to take into account such minor differences as that just reported in editions of Ammianus Marcellinus. If, for example, we run a comparison of means test to see if we are justified with 95% certainty in rejecting the hypothesis that Gardthausen's text could be a sample from the same work as Galletier's, the result is negative. That is, we are not justified in supposing that the two editors are punctuating a different original text.11

To return to Marriott, thus far in our critique, we have been speaking of philological significance. What about statistical significance? Here, too, Marriott's study claims more than we can grant. Marriott reports that he ran a difference of means test on the six HA authors. This test yields values ranging between 0% probability and 100% probability in favor the so-called null-hypothesis. Normally, one accepts the null-hypothesis when the probability is 5% or more in its favor; if the probability falls below 5%, then one rejects the null hypothesis and accepts the alternative hypothesis. In Marriott's application of the test, the null hypothesis is that the mean "sentence"-lengths of each HA author are so similar that they are probably samples of a single author. The alternative hypothesis is that the mean "sentence"-lengths for each of the HA authors are so different that they are unlikely to be samples drawn from the works of a single author.

---

10 In the edition of A. C. Clark (1918), there are 156 "sentences" with a mean of 10.94 words and a standard deviation of 7.99 words. In the edition of Shackleton Bailey (1985) the same speech is divided into 141 "sentences" with a mean of 11.16 words and a standard deviation of 8.439 words.

11 For Gardthausen (1874) the data is: Y = 22.73; Standard Deviation = 12.82; N = 170. For Galletier (1968), the data is: Y = 23.40; Standard Deviation = 12.45; N = 187. In the z-test, H0 is that Galletier's mean equals Gardthausen's (implying that the two texts are samples from the same underlying text, i.e., Books 14–16 as written and punctuated by Ammianus himself). H1 is that Galletier's mean does not equal Gardthausen's (implying that the two texts are not samples from the same underlying text). Z = .768, which has a P-value of .2236, so we fail to reject H0 at a = .05.
Marriott does not give the calculations\textsuperscript{12} but merely asserts (67–68), “if this test is applied ... the Historia Augusta is homogeneous. The six ‘authors’ show no significant difference among themselves, while they are totally distinct from any of the control texts.” Yet, \textit{pace} Marriott, if the means of the two extreme authors—Gallicanus and Pollio— are compared with this test, we find that the probability in favor of the null hypothesis is far less than 1%, and so we firmly reject the hypothesis that Gallicanus and Pollio are really the same author and accept the alternative hypothesis of their heterogeneity.\textsuperscript{13} Moreover, here, again, we find a design flaw in Marriott’s experiment. If one suspects that the six authors of the manuscripts are fictions, and that one author really wrote the entire collection, then the proper course of action is to compare—not each of the six authors against the other five—but each of the thirty biographies against the other twenty-nine. This we will presently do.

Once again, having the right control texts helps us to see the futility of using words per strong stop as an indicator of authorship. If we run a difference of means test for the overall mean of the HA as against the mean of Cornelius Nepos, we get a P value of 29.7%. This P-value is far greater than 5%, the highest value that would permit us to reject the null hypothesis that the texts are random samples from the same author. When we run the same test comparing Aurelius Victor and the HA, our P-value is 31.3%, so that once again we fail to reject the null hypothesis of unitary authorship.\textsuperscript{14} Thus, by Marriott’s reasoning, the HA and the biographies of Cornelius Nepos and of Aurelius Victor could have been written by the same person. I would argue that a more reasonable explanation would be that words per strong stop is not a good indicator of authorship in the genre of Latin biography.

The statistical test that Marriott uses is called “bivariate” and is not as efficient for the kinds of multiple comparisons we would like simultaneously to make as is a multivariate test like the chi-square test. In this test, we can compare all thirty lives of the HA at the same time, trying to determine if they have as many words per strong stop as we might expect if they were all written by the same author. Like the difference of means test, the chi-square test also yields values ranging between 0% probability and 100% probability in favor the null-hypothesis. Once again, we reject the null-hypothesis when the probability in its favor is less than 5%. In our case, the null hypothesis is that all the lives were written by the same author. The test of this hypothesis is firmly rejected, because the probability of its being correct is less than

\textsuperscript{12} And (at p. 67) he errs in giving the formula, where instead of the equal sign (=) read \(+/-\).  
\textsuperscript{13} The exact P-value is just 0.0019, meaning that the odds are only about 2 out of 1,000 that the texts attributed to Gallicanus and Pollio are actually samples of a single author.  
\textsuperscript{14} This is not to say that the difference of means test does not sometimes work: it does, for example, in the case of Suetonius vs. HA, where the P-value is 0.001. My point is not that the test never works, but that it is too unreliable to be of use in a case like the present one where authors are striving to conform to a stylistic norm.
1%: only 0.5%, to be exact.\textsuperscript{15} This is not surprising: although the chi-square test does strongly suggest that Nepos’ biographies were all written by the same person,\textsuperscript{16} the probability in favor of Tacitus’ historical works being written by the same author is only about 3.5%; and for Livy, the odds are even lower — ca. 2.0%.\textsuperscript{17} So, here, too, we find that the mean number of words per strong stop is not a very reliable indicator of authorship. The reasons are that even though the mean values in these authors are philologically close as a group, within each author they differ enough from work to work, and there are a sufficient number of cases to give these differences statistical significance.

A still more powerful test is the analysis of variance (ANOVA). Thus far, in both our test of the difference of means and of chi-square, we have considered only the mean number of words per strong stop and the standard deviation from the mean in the \textit{HA} and in our control texts. In so doing, we have been ignoring the information we have about the precise shape of the distribution of “sentence”-types: the percentage of sentences of one word, two words, three words, etc. If the number of words per strong stop is in fact a good way of discriminating one author from another, we might well expect that not only the mean “sentence” length but also the distribution of all “sentence” lengths in texts by an author provide telltale signs of his unique style. However, this turns out not to be the case, driving yet another nail into the coffin of Marriott’s hypothesis.

To show this, Nepos’ longer biographies were studied because they had done so well on the chi-square test. If we run an ANOVA test of the 429 “sentences” of Nepos’s lives of Alcibiades, Atticus, and Dion, we find that the probability is 0.00% that the distribution of “sentence”-types in each life is so similar that we may view them as samples from the same population.\textsuperscript{18} The same result (0.00%) was found when an ANOVA test of lives of Suetonius was run.\textsuperscript{19}

We may conclude, then, by observing that “sentence”-length is not a reliable indicator of authorship in our group of Roman historians and biographers. Mean “sentence”-lengths in works by authors writing in these related genres are too similar and the distributions of “sentence”-lengths in individual works by an author are too disparate to allow “sentence”-length mean or variability to help us discriminate the works of one author in the group from another. This does not necessarily mean that

\textsuperscript{15} Chi-square = 91.932; degrees of freedom are 58; \textit{P} = 0.005.
\textsuperscript{16} Chi-square = 2.38 with 12 degrees of freedom, giving a \textit{P}-value of ca. .999.
\textsuperscript{17} For Livy, chi-square = 23.31 with 12 degrees of freedom. For Tacitus, chi-square = 29.53, with 14 degrees of freedom.
\textsuperscript{18} In the ANOVA test run, the sum-of-squares was 913.026 with two degrees of freedom; the mean-square was 156.513. The F-Ratio was 34.887, which has a \textit{P}-value of 0.000.
\textsuperscript{19} In the ANOVA test, the lives of Caesar, Augustus, Tiberius, and Claudius were studied. The sum-of-squares was 2,452.74 with three degrees of freedom; the mean-square was 817.58. The F-Ratio was 6.12, which has a \textit{P}-value of 0.000.
the Dessau thesis is wrong — only that it cannot be proven, as Marriott supposed, by
using words per strong stop as a discriminator.

2. Theory and practice of “Sentence”-length in Cicero

This negative conclusion does, however, give rise to a more positive and important in-
sight into the stylistic characteristics of these genres of Latin literature. The discovery
that mean “sentence”-length of the works in the HA and our control texts fall into a ra-
ther narrow band of ca. 14–17 words (with philological, if not statistical, significance)
is unexpected and therefore interesting. How are we to account for this fact?

One place to start is rhetorical theory. Beginning with Aristotle, it is clear that the
ancients had a theory of the proper size of a period. In the Rhetoric, Aristotle said
that the orator’s period should be εὐσύνοστος ("of a size to be taken in at one view")
and εὐοναυφεύστος. This Aristotelian hapax means either “easy to repeat in a breath”
or, as Adamik has recently suggested, “easily utterable from the point of view of
breathing.” In terms of constituent parts, the ideal Aristotelian period consisted of
one or two cola. Demetrius implies an orator’s normal period-length would be two to
four cola long (16); he also recognizes the one-colon period and sets four cola as the
upper limit (17).

On the Roman side, Cicero felt that the average sentence should have four cola,
making it about as long as a line of dactylic hexameter; but he is not insistent
(Orat. 221–222). Quintilian (9.4.68) assumes that a period should be pronounced in
one breath — in terms of length, four senarit; measured in cola, he sets two as the
minimum length and four as the normal maximum, but he grants (9.4.125) that some
periods have more than four cola.22

The rhetoricians also recognized that the styles of periods should vary by literary
genre. Demetrius distinguishes between the genres of narration, dialogue, and orato-
ry. In Quintilian, there is a similar scheme, with oratory and history (9.4.15–18)
contrasted to the looser genres of the letter and dialogue (ibid. 19–20). Cicero may
well have agreed — he recognizes the genres of philosophic writing and of history

---

20 As far as I am aware, the only anticipation of this finding is in Charpin’s observation of the
number and size of sequences of phrases (measured in terms of words) in sample texts
ranging from Plautus to St. Augustine; see Charpin 1977: 222. There are, however, many
problems with Charpin’s work, including small sample size; paucity of authors in the study;
ignoring of the genre of the works; and calculation based on sentence-length, not on words
per strong stop. Nevertheless, Charpin’s thesis contains many suggestive findings that de-
serve further study that would be much more easily done today in view of advances in com-
puting and the digitization of the bulk of the Latin authors by the Packard Humanities Insti-
tute.

21 Arist. Rhet. 1409b 13ff., on which see Adamik 1984: 187–188.

22 On the subject of period length in the rhetoricians, see Schenkeveld 1964: 33.

23 19; see Schenkeveld 1964: 40.
next to oratory in the *Orator* — but in his discussion of the period in that work, he is primarily concerned with oratory. He does, however, tell us that the historian’s period in his *contiones et hortationes* is “carried along and smooth-flowing” (*tracta quaedam et fluens*, Orat. 66), while that of the orator is “vigorously and forcefully” (*contorta et acris*).\(^\text{26}\)

Demetrius makes a somewhat similar contrast between the simplicity (άπλοτης, 19) of the historian’s period as against the tenseness and roundness of the orator’s (συνεστραμμένον ... καὶ κυκλικόν, 20).

Calboli has recently suggested that with the help of the computer we are in a better position to gauge the degree to which ancient prose reflects rhetorical theory.\(^\text{27}\) The present study is a case in point. Measuring sentences in terms of words per strong stop is very different from calculating, as the ancients did, by the number of cola in a period. Words can have one, two, or more syllables and so are not a fixed unit of measurement. The colon can have few or many words; it is therefore an even more elastic measuring rod than the word. Moreover, not all Roman prose is periodic. For us, the importance of ancient theories of the period is that they existed at all and had a quantitative as well as a qualitative aspect. This encourages us to investigate the possibility that a measurement in terms of words per strong stop might turn up similarities between authors writing in related genres. Thus far, we have seen such similarities in the case of biographers and historians. Let us see if their mean “sentence”-length differs from that of the orators, as is suggested by Ciceronian theory.

Who better to study than Cicero himself? Table 4 gives the data follows for slightly over half of his 58 extant speeches: Cicero’s mean number of words per strong stop (16.55) is well within the range we have seen for the biographers and historians. However, in Cicero’s case, there is a great deal more variation around the mean, and the fluctuations may be associated with chronology. As figure 1 shows, in the earliest preserved orations and in the latest, the average number of words per strong stop is far below the overall average in the corpus as a whole. In the big middle period of Cicero’s oratorical career (say, from 70 to 52 B.C.), the number of words found in the typical clause exceeds the overall average. Thus, Cicero’s development as a speaker can be characterized as parabolic, rising from a low mean in an early speech like the *Pro Quinctio* (12.81) to a high of 23.44 in the mid-career *Post Reditum ad Populum* and then falling back into the 11-15 range in two of the three “Caesarian speeches”\(^\text{24}\) 62–66. Cicero also makes a strong distinction between Thucydides’ style and that of the orator in *De Optimo Genere Oratorum* 15.

\(^{25}\) I take *tracta* closely with *fluens* in the sense of carrying along currents of water (cf. OLD s.v. *traho* 3).

\(^{26}\) Cf. Orat. 66: *huic generi historia finitima est. in qua et narratur ornate et regio saepe aut pugna descriptur; interponuntur etiam contiones et hortationes. sed in his tracta quaedam et fluens expeditur, non haec contorta et acris oratio.*

\(^{27}\) In Norden 1988: II 1139.
and in the *Philippicae*. Given the relationship between strong stops and natural pauses for breathing, this development may well reflect the gain in strength and lung-power that Cicero tells us occurred when he studied with the rhetor, Molon, in the years 79–77. The drop in words per strong stop from 46 to the end of his life may likewise be a reflection of the fact that Cicero's oratorical powers were on the decline after what he called (Marc. 1) the *diuturnum silentium* of the years from 51 to 47 in which he gave no public speeches. He had apparently allowed himself to get out of shape, but his breath control was improving as he reached the late Philippics.

![Graph showing words per Strong Stop in 30 Ciceronian orations](image)

**Figure 1:** Graph showing words per Strong Stop in 30 Ciceronian orations

[Y axis = Mean number of words per strong stop; X axis = Date (not to scale)]

---

28 Here we may note that the anomalous *Pro Marcello* (whose mean “sentence” length is 18.64 words) is much shorter than the other “Caesarian speeches.” Whereas the *Pro Marcello* has 151 strong stops, the *Pro Ligario* has 251 and the *Pro Rege Deiotaro* 280. The source of these and the other Ciceronian texts was the CD-ROM published by the Packard Humanities Institute.

29 Cf. Brut. 312–316, especially 316: *ita recepi me biennio post non modo exercitator sed prope mutatus. nam et contentio nimia vocis resederat et quasi deferverat oratio lateribus-que vires et corpori mediocris habitus accenserat.* Of course, Cicero's published speeches do not have to correspond exactly to the original orally-delivered versions (cf. Humbert 1925), but Narducci (1991: 33) is certainly correct in calling the relationship between the lost oral originals and the preserved written versions a *vetusta, e secondo me insolubile, questione.* In lieu of any definitive reason not to do so, I will use as a working hypothesis the assumption that the preserved speeches do give us fairly accurate evidence of the “sentence”-lengths of the oral originals.

30 Cicero speaks about the orator's need to stay in good physical condition at De Orat. 3.220. A similar development in the case of Demosthenes was postulated by Pearson 1975. Pearson did not study the data for mean words per strong stop but looked only at sentences he considered particularly difficult to pronounce. If, however, we take his two extreme cases — the Against Leptines, Demosthenes' first public oration, and the Third Philippic, which Pearson called filled “with the most remarkable examples of virtuosity” (p. 227) — we find virtually no difference in the overall demands that the speeches make upon the orator. The mean number of words per strong stop in the Against Leptines is 19.13; in the Third Philippic, 19.17. The standard deviation in the Against Leptines is 14.28; in the Third Philippic, 14.24.
Breath control cannot furnish the entire explanation of the variations we find in Ciceronian practice, because our parabola is anything but smooth. In most periods of Cicero’s life, we find speeches with quite different mean “sentence”-lengths. Recent work by Adamik would suggest that such fluctuations are related to the three levels of style (as discussed by Cicero in the Orator), with longer units being characteristic of the grand style and shorter ones of the plain style.  

Thus, at Orat. 102, Cicero uses the Pro Rabirio to exemplify the high style and the Pro Caecina for the low. From the table, we see that the Pro Caecina of 69 B.C. had 16.45 words per strong stop, on average, with a standard deviation of 13.03. Adding the Pro Rabirio of 63 B.C. to

---

**Table 4: Mean words per Strong Stop in some orations of Cicero**

<table>
<thead>
<tr>
<th>Speech</th>
<th>Type</th>
<th>Date</th>
<th>Mean</th>
<th>SD</th>
<th>SD/ME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro Quinctio</td>
<td>Forensic</td>
<td>81</td>
<td>12.81</td>
<td>11.19</td>
<td>.87</td>
</tr>
<tr>
<td>Pro Roscio Amer.</td>
<td>Forensic</td>
<td>80</td>
<td>14.19</td>
<td>11.44</td>
<td>.81</td>
</tr>
<tr>
<td>Pro Tullio</td>
<td>Forensic</td>
<td>75 (ca.)</td>
<td>14.10</td>
<td>11.88</td>
<td>.84</td>
</tr>
<tr>
<td>In Caecilium</td>
<td>Forensic</td>
<td>70</td>
<td>17.17</td>
<td>11.92</td>
<td>.69</td>
</tr>
<tr>
<td>Verrines I</td>
<td>Forensic</td>
<td>70</td>
<td>18.10</td>
<td>12.46</td>
<td>.68</td>
</tr>
<tr>
<td>Pro Fonteio</td>
<td>Forensic</td>
<td>69</td>
<td>18.86</td>
<td>14.37</td>
<td>.76</td>
</tr>
<tr>
<td>Pro Caecina</td>
<td>Forensic</td>
<td>69</td>
<td>16.45</td>
<td>13.03</td>
<td>.79</td>
</tr>
<tr>
<td>Pro Cluentio</td>
<td>Forensic</td>
<td>66</td>
<td>16.54</td>
<td>11.31</td>
<td>.68</td>
</tr>
<tr>
<td>In Catilinam IV</td>
<td>Political</td>
<td>63</td>
<td>20.76</td>
<td>14.43</td>
<td>.70</td>
</tr>
<tr>
<td>Pro Murena</td>
<td>Forensic</td>
<td>63</td>
<td>16.27</td>
<td>11.87</td>
<td>.73</td>
</tr>
<tr>
<td>Pro Archia</td>
<td>Forensic</td>
<td>62</td>
<td>22.27</td>
<td>15.81</td>
<td>.71</td>
</tr>
<tr>
<td>Post Red. in Sen.</td>
<td>Political</td>
<td>57</td>
<td>22.57</td>
<td>13.51</td>
<td>.60</td>
</tr>
<tr>
<td>Post Red. ad Pop.</td>
<td>Political</td>
<td>57</td>
<td>23.44</td>
<td>15.09</td>
<td>.64</td>
</tr>
<tr>
<td>De Domino Sua</td>
<td>Forensic</td>
<td>57</td>
<td>18.06</td>
<td>14.81</td>
<td>.82</td>
</tr>
<tr>
<td>In Vatinium</td>
<td>Forensic</td>
<td>56</td>
<td>23.09</td>
<td>16.88</td>
<td>.73</td>
</tr>
<tr>
<td>De Prov. Cons.</td>
<td>Political</td>
<td>56</td>
<td>18.91</td>
<td>13.77</td>
<td>.73</td>
</tr>
<tr>
<td>Pro Balbo</td>
<td>Forensic</td>
<td>56</td>
<td>18.22</td>
<td>14.36</td>
<td>.79</td>
</tr>
<tr>
<td>In Pisonem</td>
<td>Political</td>
<td>55</td>
<td>17.28</td>
<td>12.78</td>
<td>.74</td>
</tr>
<tr>
<td>Pro Cn. Plancio</td>
<td>Forensic</td>
<td>54</td>
<td>16.83</td>
<td>12.28</td>
<td>.73</td>
</tr>
<tr>
<td>Pro Milone</td>
<td>Forensic</td>
<td>52</td>
<td>16.71</td>
<td>13.81</td>
<td>.83</td>
</tr>
<tr>
<td>Pro Marcello</td>
<td>Political</td>
<td>46</td>
<td>18.64</td>
<td>11.96</td>
<td>.64</td>
</tr>
<tr>
<td>Pro Ligario</td>
<td>Forensic</td>
<td>46</td>
<td>13.01</td>
<td>09.53</td>
<td>.73</td>
</tr>
<tr>
<td>Pro Rege Deiot.</td>
<td>Forensic</td>
<td>45</td>
<td>15.43</td>
<td>10.68</td>
<td>.69</td>
</tr>
<tr>
<td>Philip. I</td>
<td>Political</td>
<td>44</td>
<td>12.62</td>
<td>09.34</td>
<td>.74</td>
</tr>
<tr>
<td>Philip. II</td>
<td>Political</td>
<td>44</td>
<td>11.45</td>
<td>08.13</td>
<td>.71</td>
</tr>
<tr>
<td>Philip. III</td>
<td>Political</td>
<td>44</td>
<td>14.43</td>
<td>10.06</td>
<td>.70</td>
</tr>
<tr>
<td>Philip. VI</td>
<td>Political</td>
<td>43</td>
<td>10.94</td>
<td>07.99</td>
<td>.73</td>
</tr>
<tr>
<td>Philip. VIII</td>
<td>Political</td>
<td>43</td>
<td>10.67</td>
<td>08.63</td>
<td>.81</td>
</tr>
<tr>
<td>Philip. XII</td>
<td>Political</td>
<td>43</td>
<td>10.88</td>
<td>07.54</td>
<td>.69</td>
</tr>
<tr>
<td>Philip. XIV</td>
<td>Political</td>
<td>43</td>
<td>15.68</td>
<td>12.63</td>
<td>.81</td>
</tr>
<tr>
<td>AVERAGE</td>
<td>—</td>
<td>—</td>
<td>16.55</td>
<td>12.12</td>
<td>.74</td>
</tr>
</tbody>
</table>

---

table 4, we find that it has a mean “sentence”-length of 20.40, with a standard deviation of 18.35. Here, then, Ciceronian theory would appear to be confirmed by his practice.

Another factor to be considered is the influence on Cicero of contemporary Roman rhetors and orators — particularly in his formative years. In this connection it is pertinent to note that in the Ad Herennium the “sentence”-lengths of the short speeches adduced to exemplify the three styles at IV.viii.12, 13, and 14 respectively, are: Grand Style — 12.25 (S.D. = 7.216); Middle Style — 14.53 (S.D. = 9.62); Plain Style — 9.36 (S.D. = 5.76; text: Marx). Here the number of words per strong stop is considerably lower than the Ciceronian average, as we found was also the case in Cicero's early orations written about the time of the Rhetorica Ad Herennium. Thus, a full model explaining figure 1 would have to take into account, at the very least, Cicero's early training, his study of breath control, and his application of the doctrine of the three levels of style.

If we compare Cicero's practice with that of the biographers and historians, we find that the differences outweigh the similarities. The main similarity, of course, is in the overall mean number of words per strong stop. As noted, Cicero's 16.55 is within the range we have found for the historians and biographers of all periods. However, works of the latter do not vary as much as do Cicero's and show no pattern of development reflecting the writer's physical condition or use of a particular level of style. Moreover, within each work, the ratio of standard deviation to mean number of words is lower in the historians and biographers than it is in Cicero: Cicero's is .74, whereas that of the historians and biographers is only .63. This means that the degree of variability in the “sentence”-lengths of Cicero's speeches is nearly 20% greater than that found in the historians and biographers. Here it is interesting to note that the Pro Rabirio has an even higher ratio — .899. This suggests that “sentence”-length and variability taken together might be good indicators of Cicero's grand style, which among other things is characterized by forcefulness and variety at Orat. 20. These terms might have a range of stylistic applications, including dilectus verborum, syntax, rhythm, as well as “sentence”-length.

In this sense, the computer does, indeed, help us begin to test and confirm the accuracy of Cicero's comment in the Orator (66) that the historians' style is more smooth-flowing and continuous and that of the orator is more vigorous and forceful, or discontinuous. Obviously, a much more detailed study would be required to test the full conformity of Cicero's practice to his theory of “sentence”-length. For example, at Orat. 66, Cicero does not contrast the historians' style generally to the orators' (as we have done), but specifically comments on the differences in their handling of

32 I have not included the fragments in the count.
speeches. He also characterizes the philosophical style, which suggests that it might be instructive to compare his philosophical works to his speeches. At any rate, the results reported here may be taken as giving strong encouragement to the potential value of further study, and they lead to the provisional conclusion that Cicero’s actual practices involving words per strong stop do, indeed, reflect his theory about the generic factors determining prose style.

REFERENCES

—. 1892. Über die SHA. *Hermes* 27, 561–605.

---

33 In qua et narratur ornate et regio saepe aut pugna describitur; interponuntur etiam contiones et horationes. Sed in his tracta quaedam et fluens expetitur, non haec contorta et acris oratio. Cicero contrasts the style of the orator with that of the historian in a general way at *De Optimo Genere Oratorum* 15.

* I wish to thank the Research Committee of the UCLA Academic Senate for its generous support of the research that went into this article. Versions of this paper were given at the UCLA-USC Latin Seminar, the UCLA Data Seminar, and the Seventh International Colloquium on Latin Linguistics. I thank the following for their helpful comments on my work: Tamás Adamik, Victor Bers, David Blank, Gualtiero Calboli, Basil Dufallo, Andrew Dyck, and Haiim and Hannah Rosén.


