QUANTITATIVE LINGUISTICS
Vol. 11

Hexameter Studies

Rüdiger Grotjahn

Studienverlag Dr. N. Brockmeyer
Bochum 1981

CIP-Kurztitelaufnahme der Deutschen Bibliothek

Grotjahn, Rüdiger:
Hexameter studies / Rüdiger Grotjahn. – Bochum:
Studienverlag Brockmeyer, 1981.
(Quantitative linguistics; Vol. 11)
ISBN 3-88339-196-4

NE: GT

ISBN 3-88339-196-4

Querenburger Höhe 281, 4630 Bochum 1
Druck Thiebes GmbH & Co. Kommanditgesellschaft Hagen

CONTENTS

INTRODUCTION

1. General Remarks

If one asks of the various disciplines included under the concepts "philology" and "linguistics" how far they use statistical methods, one finds that among these metrics has a relatively long tradition of applying statistical procedures. Probably the principal reason for this is that the analysis of versified language is in the first place an analysis of the *form* of language and text, a form characterized on the level of concretion by a large amount of variability. The consequence of this is that only by using statistical procedures is it possible to carry out analyses in the field of metrics in a scientifically adequate way. Such a statistically-oriented metrics, which in analogy to the term "quantitative linguistics" may be called "quantitative metrics", not only attempts to cover manifest characteristics of metrical systems and metrical texts, but also tendencies hidden under the surface.

The use of statistical methods for the analysis of verse, and in particular for the analysis of classical hexameter, begins as early as the middle of the previous century, and as early as the 1900s Russian Formalism began systematically to develop a statistically-based metrics (a development which continued within the framework of Czech and Polish structuralism). However, we are still far from having any statistically-based theory of metrics or even of any single metrical system. This statement even applies (although in a somewhat weaker form) to individual verse types such as hexameter. This is of course not surprizing when one considers how many metres there are, how many languages and cultures, and by how many variables (e.g. the structure of the language, the poetic canon, individual style, etc.) the metrical system and in particular the versified text in question is determined.

Inside this large number of metrical systems, classical metrics has a special importance, particularly classical hexameter because of its decisive influence on the formation and development of the metrical systems of "modern" European literature. For instance, for many years the value of a national metrical system was measured by its ability to recreate the "sound" of classical verse.

The contributions to this volume extend our knowledge of hexameter with the help of often very extensive corpora (primarily from Greek and Latin authors) with subsequent statistical analyses. A no less important goal of this volume is to demonstrate applications of statistical procedures to metrics, and in particular to research on hexameter. The intention of this is to create a more objective basis for metrics, which to date has very often been concerned with the analysis of isolated data, and exclusively oriented towards a philological-interpretative method. At the same time we hope that researchers in the field of metrics who do not yet make use of the methods of statistics may be convinced of the advantages of these procedures.

2. Contents

Considering the vast number of different potential problems in the analysis of hexameter, it is clearly only possible in this volume to deal with a very limited catalogue of topics; these can be subsumed under the following partially interdependent areas:

1. The analysis of the mutual influence of metrical structures and the phonological/syntactical/lexical system of the particular language in question, with regard to their dependence on factors such as individual style, period, poetical canon, etc.
2. The analysis of the recurrence of metrical elements (e.g. verse feet, caesurae in particular authors, groups of authors, or with different national hexameter types.

SCHMIEL's article "Rhythm and Accent: Texture in Greek Epic Poetry", which can be considered in the first category, is primarily devoted to the study of the relation of ictus and accent in Greek epic poetry. To this end the frequencies of coincidence of ictus and accent are established on the basis of extensive samples taken primarily from the *Iliad* and the *Odyssey*. Subsequently the hypergeometric probability distribution is used to calculate the probability of obtaining 0,1,... coincidences in lines consisting of N syllables out of which n are accented and k are heavy. The probabilities obtained are then used to classify lines as homodyne, heterodyne or neutral. Analysing the relationship between form (homodyne, heterodyne, neutral) and content (narrative, speech, calm etc.) with the help of χ^2 tests, SCHMIEL, for example, comes to the conclusion that vivid lines tend to be homodyne whereas calm lines tend to be heterodyne.

In his article "A Statistical Model for the Analysis of the Coincidence of Ictus and Accent", GROTJAHN discusses the hypergeometric probability distribution as a general model for the analysis of the coincidence of ictus and accent. By means of some simple combinatorial deliberations, the hypergeometric probability distribution is derived, and some characteristics of this distribution are given which are useful for practical applications. After a critical analysis of the classification attempt by SCHMIEL, a table is presented permitting an objective classification of hexameter lines into homodyne, heterodyne and neutral. Finally, the use of some approximations to the hypergeometric model in the analysis of longer passages of verse is discussed and a coefficient is presented which allows one to measure the strength of association between ictus and accent.

TORDEUR in his "Etude statistique sur l'hexamètre d'Ausone" attempts to determine the style of several poems by Ausonius according to the frequency of the different metrical word types. The frequencies are first analysed with the help of χ^2 tests. The χ^2 analysis is then supplemented by a statistical test which allows him to determine those cells of his extensive contingency tables which show significant discrepancies between observed and expected frequencies. The statistical analyses lead TORDEUR to the conclusion that there are clear differences among certain of the poems. These differences seem to be relatively large, in particular for instance between the poems categorized as "négligé" and those categorized as "soigné" as well as between *Technopaegnion* and the remaining poems.

In his "Note on Pol Tordeur's Paper 'Etude statistique sur l'hexamètre d'Ausone'", GROTJAHN first comments on the test for deviation in a single cell used by TORDEUR. On the basis of TORDEUR's data GROTJAHN then tries to show that without the computation of appropriate coefficients which measure the relation between two or more variables independently from the sampling error, one may arrive at conclusions in the analysis of contingency tables which may be not completely justified.

In his article on "Sentence Length in Greek Hexameter Poetry" CLAYMAN analyses a sample consisting of 23,049 sentences gathered from all hexameter poetry from the eigth to the second centuries B.C. He finds that when sentence length is measured by word, the sentence length distributions in Greek hexameter poetry are similar to those found in Greek prose. When sentence length is, however, measured by phonemes (or by syllables), multipeaked distributions are produced which show that in hexameter poetry the end of the sentence tends to coincide with the end of the verse. Although this tendency seems to be strongest in earlier poems, it is argued that exceptions counterindicate the use of sentence length distributions to date a poem or to decide whether it was composed orally. Analysing the homogeneity of individual poems with regard to sentence length distributions, CLAYMAN comes to the conclusion that there are subtle variations in sentence length distributions measured by phoneme which in many cases can be related to changes in subject and matter.

ALTMANN's article "The Homogeneity of Metric Patterns in Hexameter" is devoted to demonstrating how to use so-called information statistics in the analysis of the frequencies of different types of hexameter lines. To this end ALTMANN distinguishes 16 different types of hexameter lines as defined by the 16 possible combinations of dactyls and spondees in the first

four feet of a line. Analysing the frequencies of the 16 types in samples from several Latin and German authors, he shows how one can decide with the help of information statistics whether one or more texts, one or more authors, or one or more groups of authors can be considered as homogeneous.

GREENBERG's article is devoted to the study of "Vocalic Initials in *Aeneid* 12". To this end GREENBERG analyses the distribution of both initials and elisions at each position in the verse line. In the course of his analyses he isolates some anomalous aspects of that distribution and attempts some explanation for these primarily on the basis of an analysis of the distribution of the different metrical word types at the positions 3, 6, 9, 12, and 15 of a line. He shows, for instance, that there are three positions within the line, namely positions 3, 6, and 9 (following the final syllables of feet 1, 2, and 3), where the proportion of vocalic initials is considerably higher than that found at the beginning of the verse. If at these three positions the preceding foot is a spondee, a very high percentage of the vocalic initials is involved in elision. However, if the preceding foot is a dactyl, the percentage of elisions is much lower. Although GREENBERG uses the *Aeneid* 12 as a paradigm case, the major finding of his study seems to be true of Latin hexameter verse in general.

In his article "Liaisons syllabiques dans les partages trochaïques de l'hexamètre latin" GÉRARD differentiates between several types both of trochaic division and syllabic liaison. On the basis of extensive samples from numerous Latin authors from Ennius to Juvenal, he establishes the frequencies of each type and analyses them with the help of some descriptive statistics. On the basis of these analyses a number of interesting conclusions are drawn, such as that both in trochaic division marked by strong punctuation and trochaic division marked by weak punctuation vowel-consonant liaisons tend to be more frequent than consonant-vowel liaisons in the category 'content word' and less frequent than consonant-vowel liaison in the category 'function word'.

GROTJAHN in his "Note on Jean Gérard's Paper 'Liaisons syllabiques dans les partages trochaïques de l'hexamètre latin'" reanalyses part of GÉRARD's data by applying some simple methods of inferential statistics. It is shown how these methods can be used to test some of GÉRARD's hypotheses as well as some additional ones. GROTJAHN's statistical analyses more or less confirm the central points of GÉRARD's conclusions and in addition show that, for example, the frequency of weak trochaic punctuations in the first foot tends to increase chronologically from Ennius to Juvenal.

JOB in her "Annotated Bibliography on the Statistical Study of Hexameter Verse" presents information on books and articles on quantitative analysis of hexameter poetry. The period under review covers the time from 1866 to 1981. All of the 249 entries (apart from a few inaccessible ones) have been supplied with a short descriptive note. The principal aim of this bibliography is to provide information on data already collected and methods already in use in the analysis of hexameter verse and thus to stimulate further research in this field.

Bochum, January 1981 R. Grotjahn

Rhythm and Accent: Texture in Greek Epic Poetry

Robert Schmiel, Calgary

The investigations of W. F. Jackson KNIGHT (1939) into the texture of Vergil's poetry have contributed greatly to the understanding and appreciation of Vergil's artistry. The purpose of this article is to apply the study of texture to Greek epic poetry, Homer in particular.

KNIGHT begins his study with the statement that all sounds of the human voice have quantity, pitch, and intensity (p. 3). He describes the Latin hexameter so:

> Greek hexameters used by Latin poets have metrical ictus. The first syllable of each foot has a certain emphasis, not easy to define, since the principle of emphasis is not exactly stress, pitch, quantity, or even loudness.... The ictus is connected with rhythm and regular recurrence. The mind expects a certain recurring moment; the voice recognizes its recurrence with a certain emphasis released after an expectant delay. The regularity of the metre therefore confers an emphasis on certain syllables. (p. 4)

KNIGHT argues that in the Latin hexameter there were two waves, an ictus wave and an accent wave (p. 10 f.). He uses the analogy of wireless telegraphy, in which one wave is superimposed upon another, in order to show the effect of coincidence or lack of coincidence between the two waves (p. 12 f.). "Homodyne [coincidence between ictus and accent] makes the line free, ready and smooth. Heterodyne [lack of coincidence] makes it constricted, reluctant, and rugged" (KNIGHT 1944: 240 f.)

In Chapter IV KNIGHT asserts that the texture is regularly expressional in Vergil, but not in other Latin poets (p. 15). He cites such lines as *Aen.* 8. 452: "îlli inter sese multa vi brácchia tóllunt", where heterodyne texture (in combination with the spondees) produces the effect of great effort; and *Ecl.* 1, 70: "împius hâec tam cûlta novália mîles habébit", where the effect is one of speed and unrestrained indignation. One might think that the difference is primarily metrical, but *Aen.* 3. 658 disproves that notion: "mónstrum horréndum infórme ingens cui lúmen adémptum". The line is as spondaic as the first example, and yet it suggests slow but irresistible movement (p. 18). Knight also analyzes longer passages such as *Aen.* 1. 1-33, 81-91, and 278-96. The comparison of the storm in *Aen.* 1 with the one in Valerius Flaccus, *Argonautica* 1, is particularly, convincing in that, despite apparent similarities, Vergil's description is much more expressional in texture (p. 26 ff.). KNIGHT's analysis of *Aen.* 1. 29-33 will serve as an example: the numbers show how many coincidences between ictus and accent there are in each line:

4 his accénsa super iactátos áequore tóto
3 Tróas, reliquias Danaum atque inmîtis Achîlli
2 arcebat longe Latio multósque per ánnos
2 errabant acti fatis maria ómnia cîrcum
5 tántae mólis erat Románam cóndere géntem.

The third and fourth of these verses express the weariness of the Trojan's long quest by the heavy reluctance of their texture..., which is further increased by the metrically similar first two words in each verse. Words composed of three long syllables are peculiarly direct, free, and forcible positively, when they are homodyned; and they are just as forcible negatively when they are not, as here. The last verse has ... a strong and triumphant finality in its five-homodyne. (p. 22)

More than half of KNIGHT's book is devoted to fourth-foot texture, with emphasis on its occurence in simple, compound, and complex patterns. Since that portion of his work has no relevance to the present study, I shall not discuss it. There has

been ample time to evaluate KNIGHT's work, and some of his conclusions have been accepted by scholars, but his views have not had much effect on practical criticism. One likely reason is that, because of his horror of statistics (p. 36), KNIGHT did not provide an objective statistical proof of his basic contention that texture is expressional. GETTY's critical bibliography summarizes the judgement of KNIGHT's work at present: "Upholders of the ictus-accent theory in the English-speaking countries seem largely to agree in principle with his feeling, although not necessarily with the more or less elaborate patterns of fourth-foot texture." (GETTY 1963: 132).

The many other similarities between Greek and Latin poetry suggest that the relationship between ictus and accent in Latin poetry might have its analogue in a relationship between metrical rhythm and accent in Greek poetry[1]. The Greek word accent will hardly have been submerged without a trace under the weight of metrical rhythm. And if the accent is there, the poet will presumably accomodate it or make use of it. Aristotle explicitly states that the use of the different accents was a poetic skill, and that this skill should be discussed in connection with metrics:

δῆλον οὖν ὅτι καὶ περὶ τὴν ῥητορικὴν ἔστι τὸ τοιοῦτον ὥσπερ καὶ περὶ τὴν ποιητικήν, ὅ περ ἕτεροί τινες ἐπραγματεύθησαν καὶ Γλαύκων ὁ Τήιος. ἔστι δὲ αὐτὴ μὲν ἐν τῇ φωνῇ πῶς αὐτῇ δεῖ χρῆσθαι πρὸς ἕκαστον πάθοσ, οἷον πότε μεγάλῃ καὶ πότε μικρᾷ καὶ πότε μέσῃ, καὶ πῶς τοῖς τόνοις, οἷον ὀξείᾳ καὶ βαρείᾳ καὶ μέσῃ.

ταῦτα δὲ (τὰ στοιχεῖα) διαφέρει σχήμασίν τε τοῦ στόματος καὶ τόποις καὶ δασύτητι καὶ ψιλότητι καὶ μήκει καὶ βραχύτητι, ἔτι δὲ ὀξύτητι καὶ βαρύτητι καὶ τῷ μέσῳ. περὶ ὧν καθ' ἕκαστον ἐν τοῖς μετρικοῖς προσήκει θεωρεῖν.

It is clear, then, that there is a kind of skill similar to that in regard to poetry, and this has been treated by Glaucon of Teos and others. It has to do with the voice; how one should use it for each emo-

> tion, as, for example, when one should make it loud and when soft and when intermediate, and how one should use the accents, namely the acute, the grave, and the middle accent.
>
> And these speech sounds differ in the configuration of the mouth, in the place of articulation, in roughness and smoothness (i.e. aspiration or the lack of it), in length and shortness, and also in the acute, grave, and middle accent; each of which topics should be discussed in connection with metrics.[2]

One might reasonably object that, granted the impossibility of reading and hearing ancient poetry precisely as the ancients did, attempts to recapture the subtle modulations of poetry gone partially mute are bound to fail. But such a counsel of despair should no more be heeded here than, for example, in the case of the pre-Socratic philosophers, fragmented, mutilated and misinterpreted beyond hope of full comprehension. Working with less than satisfactory evidence or data is business as usual for the classicist. Here the task is to show that traces of the living poetry exist in Homer's use of the accent even as they exist in the comprehensible realities of Greek meter. MAAS' proposal that we give up the attempt to pronounce Greek correctly and get on with the study of poetry is quite wrongheaded. Without sound "the chief purpose of study, which is to understand the texts" is not possible; for poetry does not exist apart from sound (MAAS 1962: 55 f.).

There are, admittedly, certain problems and obstacles in the way of this investigation. The nature of the Greek (pitch) accent, the grave in particular, is not fully understood. "Ictus" is no longer a useful term and must be discarded. We may begin with this description of the dactylic hexameter and its movement: The dactylic hexameter consists of six regularly-recurring heavy syllables separated from one another by two morae (one heavy or two light syllables), the last heavy being followed by a single syllable which signals the end of the line. These six regularly-recurring heavy syllables articulate the line and produce a pulse or rhythm which is expected line after line.[3] Because these six syllables are heavy and expected, they

are picked out or emphasized. This is a normative phenomenon, and that is why the erratically-occurring heavy syllables do not have emphasis; they are not expected.

One can not readily describe how this effect is produced or perceived, whether as "stress, pitch, quantity, or even loudness" (KNIGHT 1939: 4).[4] As WILKINSON (1966: 95) says: "Beare is ... surely right when he insists that ictus was not expiratory: It is not something heard or seen: it is expectation. Expectation of a stressed syllable in accentual verse, of a heavy syllable in quantitative verse." But how the effect was produced or realized is not vital to the present study if the locations of expectation or emphasis are known. The two properties, heavy quantity and regularity of occurence, are indisputably real.

As I scanned many hundreds of lines of Homer, it seemed to me that speech was more homodyne than narrative. This observation was checked in the following way: The criterion for homodyne was arbitrarily set at 60% or better. That is, *any line with 60% or more coincidence between the six regularly-recurring heavy syllables and acute and circumflex accents (based on the number of acutes and circumflexes) was considered homodyne.* All the rest was considered non-homodyne. The results of the analysis of 12 books of Homer were: speech 29.1% homodyne and narrative 25.8% homodyne. The difference does not appear obviously significant. But a chi square analysis showed significance at the .010 level. (Chi square: 9.217; $X^2_{(1;0.01)} = 6.635$)[5] [See Table 1].

Since there is some question regarding the grave accent, whether it indicates a rise in pitch or a fall in pitch or less rise than the acute, statistics were also collected taking the grave into account along with the acute and circumflex. The results, using the sample as before, were quite different [see Table 2]: speech 19.0% homodyne and narrative 17.5% homodyne. The percentages are much lower, and the difference is only 1.5%. The chi square analysis now fails to show significance even at the .050 level, and is indeed well below that level (Chi square

Table 1. Sample BA of homodyne lines in Homer

Book	Type	Lines n_i	Homodyne n_i	%	n_i	%	Chi square
Il.1	speech	373	120	32.2	[123	33.0]	1.626
	narr.	238	65	27.3	[64	26.9]	[2.533]
Il.3	speech	242	66	27.3	[66	27.3]	1.204
	narr.	219	50	22.8	[47	21.5]	[2.098]
Il.4	speech	242	60	24.8	[59	24.4]	.000
	narr.	302	75	24.8	[74	24.5]	[.001]
Il.9	speech	584	171	29.3	[160	27.4]	5.247
	narr.	125	24	19.2	[22	17.6]	[5.180]
Il.10	speech	292	87	29.8	[90	30.8]	.626
	narr.	287	77	26.8	[75	26.1]	[1.562]
Il.21	speech	270	76	28.1	[80	29.6]	.307
	narr.	341	103	30.2	[103	30.2]	[.024]
Od.4	speech	617	174	28.2	[174	28.2]	7.943
	narr.	230	43	18.7	[39	17.0]	[11.254]
Od.7	speech	194	62	32.0	[59	30.4]	.143
	narr.	153	46	30.1	[45	29.4]	[.041]
Od.9	speech	141	49	34.8	[50	35.5]	1.057
	narr.	425	128	30.1	[121	28.5]	[2.454]
Od.15	speech	348	109	31.1	[109	31.1]	1.575
	narr.	209	55	26.3	[57	27.3]	[1.023]
Od.22	speech	208	60	28.8	[63	30.3]	4.278
	narr.	293	61	20.8	[58	19.8]	[7.312]
Od.23	speech	224	54	24.1	[56	25.0]	.038
	narr.	148	37	25.0	[41	27.7]	[.338]
Total	speech	3735	1088	29.1	[1089	29.2]	9.217
	narr.	2970	766	25.8	[748	25.2]	[13.118]

Table 2. Sample BB of homodyne lines in Homer

Book	Type	Lines n_i	Homodyne n_i	%	Chi square
Il.1	speech	373	68	18.2	.044
	narr.	238	45	18.9	
Il.3	speech	242	49	20.2	2.975
	narr.	219	31	14.2	
Il.4	speech	242	37	15.3	.054
	narr.	302	44	14.6	
Il.9	speech	584	106	18.2	1.003
	narr.	125	18	14.4	
Il.10	speech	292	48	16.4	.181
	narr.	287	51	17.8	
Il.21	speech	270	48	17.8	1.666
	narr.	341	75	22.0	
Od.4	speech	617	119	19.3	.848
	narr.	230	38	16.5	
Od.7	speech	194	44	22.7	.061
	narr.	153	33	21.6	
Od.9	speech	141	32	22.7	1.426
	narr.	425	77	18.1	
Od.15	speech	348	72	20.7	.337
	narr.	209	39	18.7	
Od.22	speech	208	48	23.1	3.145
	narr.	293	49	16.7	
Od.23	speech	224	37	16.5	1.341
	narr.	148	18	12.2	
Total	speech	3735	708	19.0	2.540
	narr.	2970	518	17.4	

Chi square analysis: 2 rows, 2 columns, 1 degree of freedom;
$\chi^2_{(1;.05)} = 3.84$; $\chi^2_{(1;.02)} = 5.41$; $\chi^2_{(1;.01)} = 6.64$

2.540; $\chi^2_{(1;0.05)} = 3.841$). [See Table 2.] Remember that the sample is the same, that most homodyne lines in the one sample are homodyne in the other, and that many lines have no grave accents and so remain precisely the same in both samples. Even if the chi square analysis of the second sample merely showed a lower degree of significance, say significance at the .050 level, the differences between the two samples would still be significant.

Following the advice of Dr. R. GROTJAHN, who pointed out a flaw in my method of determining texture, i.e. the failure to take into account the number of syllables per line (cf. in this volume R. GROTJAHN, A Statistical Model for the Analysis of the Coincidence of Ictus and Accent), I made use of tables of the hypergeometric probability distribution by LIEBERMAN and OWEN (1961). The nomenclature is as follows:

N = the number of syllables in the line
n = the six regularly-recurring heavy syllables
k = the number of unaccented syllables in the line
x = the number of unaccented regularly-recurring heavy syllables.

So, to refer ahead to the examples from *Il.* 1. 237-41, line 237 has 15 syllables (N), 6 regularly-recurring heavy syllables (n), 8 unaccented syllables (k), and 2 unaccented regularly-recurring heavy syllables (x). One need only look up N = 15, n = 8, k = 6, and x = 2 in the tables to find the result P(x) = .230769. Note that, when k is greater than n, the two values must be interchanged. I shall give results rounded off to the third decimal point, i.e. .231 for the example given. Since the upper limit for homodyne was arbitrarily set at .300, this line is homodyne. (See the tables taken from LIEBERMANN/OWEN at the end of this paper).[6]

As an example I shall apply the various methods to *Il.*1. 237-41. Coincidence between regularly-recurring heavy syllables and acute or circumflex accents (Sample BA) is marked with an x, coincidence between regularly-recurring heavy syllables and

grave accents is marked with a ´. The first pair of numbers after each line does not count the grave, the second does. The number behind the slash gives the number of accents. The number of syllables is given in parentheses.

237	φύλλα τε καὶ φλοιόν· νῦν αὖτέ μιν υἷες Ἀχαιῶν	(15)	4/7 5/8
238	ἐν παλάμης φορέουσι δικασπόλοι, οἵ τε θέμιστας	(17)	1/5 1/5
239	πρὸς Διὸς εἰρύαται· ὁ δέ τοι μέγας ἔσσεται ὅρκος·	(17)	2/5 3/7
240	ἦ ποτ᾽ Ἀχιλλῆος ποθὴ ἵξεται υἷας Ἀχαιῶν	(16)	3/5 3/6
241	σύμπαντας· τότε δ᾽ οὔ τι δυνήσεαι ἀχνύμενός περ	(16)	4/6 4/6

By the first method (homodyne = 60% or more coincidence based on the number of acute and circumflex accents) lines 240 and 241 are homodyne; by the second method (as above, but taking the grave accent into account) lines 237 and 241 are homodyne; by the third method (using the LIEBERMANN / OWEN tables and taking the acute and circumflex accents into account) lines 237, 240, and 241 are homodyne.

We may pause for a moment to consider the import of these results. Since there is a significantly higher percentage of homodyne lines in speech than in narrative (using acute and circumflex), it appears that Homeric speech and Homeric narrative differ in texture. But when the grave is used along with the acute and circumflex the difference almost disappears. Since there is doubt about the grave accent (evidence, however, suggests that the grave indicates at least a lowering of pitch from the acute), it is more reasonable to use the results based on the known than the results based partly on the known, partly on the unknown. The message transmitted by a clear signal (acute and circumflex) is preferable to the message transmitted by a partly unclear signal (for the grave may be, and probably is, static). If one accepts the former results (Sample

BA) the latter can be explained easily - the grave does not indicate a high pitch, or as high a pitch as acute and circumflex; but if one accepts the latter results (Sample BB), there is no convenient means of explaining away the former. That is, if acute, circumflex, and grave all indicate a phonetically comparable high pitch, why do Homeric speech and Homeric narrative show a significant difference in texture when the acute and circumflex are taken into account, but not when the grave is added? Where should one begin to look for an explanation of so perversely irrational a phenomenon? We must conclude, at least as an hypothesis, that Homeric speech and Homeric narrative differ in texture; and that texture in Greek poetry can be studied by using the acute and circumflex accents. We shall see if the study can be refined, and if further results will support the hypothesis.

One might wonder why Homeric speech and narrative differ in texture. I suggest the following. The coincidence of accent with regularly-recurring heavy syllables produces a stronger rhythm, whereas the accent would tend to hold the line back if it fell out of phase; a strong pulse is appropriate to a vivid mode of expression; and speech is by nature a more vivid and dramatic mode than narrative, if one will grant the difference between merely reporting an event and actually presenting the event as it happens.

This explanation raises another question. If the real distinction is between what is vivid or dramatic and what is not, then the distinction between speech and narrative is very crude and not wholly to the point. No one could reasonably maintain that speech is always very dramatic and narrative always calm. And if it should prove possible to make use of this distinction, then the statistical result should be even more decisive.

Four books of Homer were analysed in the following way. All lines were sorted according to the three types: homodyne, neutral, and heterodyne. Homodyne was defined as before: no higher than .300 on the Liebermann & Owen tables. The simplest method for determining heterodyne texture using hypergeometric

distribution was to change the definition of defective items from unaccented to accented syllables. Then the values could be read directly from the tables. So, to refer to the examples on p. 9, the P(x) value for *Il*.1. 238 would be determined as follows: the line has 17 syllables (N), 6 regularly-recurring heavy syllables (n), 5 accented syllables (k), and 1 accented regularly-recurring heavy syllable (x). One need only look up N = 17, n = 6, k = 5, and x = 1 to find P(x) = .394635. Since heterodyne was arbitrarily defined as having a value no greater than .700 on the Liebermann & Owen tables, the line is considered heterodyne. (Lines which are neither homodyne nor heterodyne are considered neutral). The speech/narrative division was similarly replaced by three categories: vivid, average, and calm lines. The terms are deliberately general and colorless, and do not, of course, make judgments of value or artistic excellence. The absolute objectivity of the speech/narrative classification had to be forfeited, for it is clearly impossible to categorize lines as vivid, average, or calm with hope of achieving complete acceptance. One should only hope that the statistical result would be so overwhelming as to cover that inevitable amount of indeterminacy. It was a matter of attempting to establish breaks in a continuum, and it proved impossible to establish fixed rules by which to sort lines. But lines with words expressing violence or deep emotion (kill, fear, etc.) were usually considered vivid, and "filler lines" (which add non-essential information to a preceding line) or lines which introduce or follow speeches were usually considered calm. A nine-line sample (*Od*. 22. 41-49) will illustrate the procedure in a general way and also show some of the problems in classification:

41	νῦν ὑμῖν καὶ πᾶσιν ὀλέθρου πείρατ' ἐφῆπται."	6/6
42	ὣς φάτο, τοὺς δ' ἄρα πάντας ὑπὸ χλωρὸν δέος εἷλε	2/5
43	πάπτηνεν δὲ ἕκαστος ὅπῃ φύγοι αἰπὺν ὄλεθρον·	1/5
44	Εὐρύμαχος δέ μιν οἶος ἀμειβόμενος προσέειπεν	1/5
45	"εἰ μὲν δὴ 'Οδυσεὺς 'Ιθακήσιος εἰλήλουθας,	1/2

46 ταῦτα μὲν αἴσιμα εἶπας, ὅσα ῥέζεσκον Ἀχαιοί, 3/6

47 πολλὰ μὲν ἐν μεγάροισιν ἀτάσθαλα, πολλὰ δ' ἐπ' ἀγροῦ. 1/3

48 ἀλλ' ὁ μὲν ἤδη κεῖται ὃς αἴτιος ἔπλετο πάντων, 5/5

49 Ἀντίνοος· οὗτος γὰρ ἐπίηλεν τάδε ἔργα, 2/5

In the comments below the lines are marked homodyne, neutral, or heterodyne as well as vivid, average, or calm (by means of the signs +, O, and -).

41 6/6 Hom. +: The line could not be more dramatic. Odysseus, who has just revealed his identity, threatens the suitors with death.

42 2/5 Neut. O: "Thus he spoke" would be considered calm, but "pale fear" is vivid. So the line is considered average.

43 1/5 Het. O: The suitors are looking for a place to flee. Though there is drama here, it is only beginning and there is no real action yet. The line may be considered average, but the texture is heterodyne. Perhaps the heterodyne texture reflects the indecision and uncertainty.

44 1/5 Het. -: The line merely introduces a speech.

45 1/2 Neut. O: The speech as yet lacks direction, but it cannot be considered calm since it refers to the (possible) return of Odysseus.

46 3/6 Neut. O: The admission of guilt is neither vivid nor calm because the point of the speech is still unclear.

47 1/3 Neut. O: The texture is neutral, and the line is considered average. It is as equivocal as the two preceding lines.

48 5/5 Hom. +: Now comes the point; the guilty man is dead.

49 2/5 Neut. O: The line is not considered vivid because it only adds to the previous line.

The most significant results are in the homodyne and heterodyne, and the vivid and calm categories (see Table 3). Nearly two-thirds of the vivid lines are homodyne, and less than 1/8th are heterodyne. Conversely, almost two-thirds of the calm lines are heterodyne, and only about one-ninth are homodyne. The chi squares range from 70.591 - 112.938, far beyond $\chi^2_{(4;0.01)}$ = 13.280. The result is clear, granted the imprecision of the operational procedure. By an overwhelming plurality, vivid lines are homodyne, not heterodyne; and calm lines are heterodyne, not homodyne. The hypothesis that degree of vividness is the real determinant of texture rather than whether the line is in the speech mode or the narrative mode is supported by the results.

Table 3. Sample CA: Homodyne, Neutral, and Heterodyne Lines in a Vivid, Average, or Calm Context

Book	Type	Homodyne	Neutral	Heterodyne	Chi square
Il. 1	vivid	67	21	6	112.938
	average	100	139	134	
	calm	20	41	83	
Il. 21	vivid	68	38	12	70.591
	average	108	142	184	
	calm	7	21	31	
Od. 22	vivid	34	11	13	70.954
	average	81	152	152	
	calm	5	9	44	
Od. 23	vivid	30	9	4	76.780
	average	65	98	113	
	calm	2	9	42	

Chi square analysis: 3 rows, 3 columns, 4 degrees of freedom

$\chi^2_{(4;.05)} = 9.49$; $\chi^2_{(4;.02)} = 11.67$; $\chi^2_{(4;.01)} = 13.28$

The evidence just presented is the foundation of my thesis, but some supplementary evidence for expressional texture can be adduced.

1) The occurrence of homodynes by metrical feet (not counting the grave). [See Table 4.]

Table 4. Occurrence of homodyne by metrical feet

	1	2	3	4	5	6
Il. 1	110	47	66	81	103	116
Il. 4	108	50	54	66	68	110
*Od.*23	108	62	76	73	95	113
Total	326	159	196	220	266	339

Each of the three samples covers the first 252 lines of the book in question. It is interesting to compare the percentages of occurrence in Homer and Vergil. The figures for Vergil are from ALLEN (1964: 8), and beneath them are the percentages from the sample above.

Table 5. Homodynes in Vergil and Homer

	1	2	3	4	5	6
Vergil	55%	28%	15%	28%	97%	99%
Homer	42%	21%	25%	30%	35%	45%

There is a very similar tendency in the distribution of homodynes in the two poets. Of greatest interest is the fact that both poets show most coincidence at the beginning and in the last two feet. The cause of the exceptionally high percentages

of coincidence in the last two feet of the Vergilian line is a celebrated problem. I am not satisfied to dismiss it as a mere byproduct of the rules for the positions of word-endings and for final-word length, especially since Homer enjoyed far greater freedom in these matters than did Vergil. I agree with those who see this as evidence of the poet's desire to keep the rhythm clear and strong at the end of the line.[7]

2) Heterodyne passages.

It soon becomes apparent when one studies the texture of Homeric poetry, that homodyne and herodyne lines do not occur in a random scattering, a fact which speaks for the purposeful use of texture. The most exciting books and passages are generally more homodyne than others, e.g. *Od.* 9. Again, some passages have a noticeably heterodyne texture. These requirements were set for heterodyne passages: they must be at least six lines long; they must be self-contained (i.e. marked by indentation in the text or by change from speech to narrative or vice versa); they must contain no more than one coincidence per three acute and circumflex accents.[8]

None of the 29 heterodyne passages that occur in the twelve books of Homer analyzed (see Table 6) was judged to be vivid (+) and only eight were considered average (O). Twenty-one were rated calm (-). The distribution by book is uneven, ranging from no examples in *Il.* 10 and 21 and Od. 9 and 15 to eight in *Il.* 4. Twenty-six occur in narrative (N), only three in speech (S), another indication of the difference between speech and narrative in Homer.

The three exceptions are instructive. At *Il.* 4. 31-49 Zeus replies to Hera who has just criticized his management of the war. In such situations Zeus and Hera usually rail at one another like fishmongers. In this instance, however, Zeus is very conciliatory, almost begging on behalf of Troy, but still giving in. Homodyne texture would be inappropriate.

At *Il.* 4. 404-410 Agamemnon has been making the rounds, trying to whip up his men with insults. He addresses Diomedes, but he refrains from replying out of respect. Sthenelos, a minor character who had not even been spoken to, replies. His

Table 6. Homodyne and Heterodyne Passages in Homer

Book	Lines	Mode	Conc.	Sense of the Passage	Context	Type
Il. 1	245-253	N	12/43	Description of Nestor about to speak	-	a
	304-311	N	11/34	Assembly ends; Agam. prepares ship	-	a
	428-411	N	12/51	Preparations for the voyage	-	a
	475-487	N	17/57	The voyage	-	b
Il. 3	111-120	N	13/50	Truce; disarming; heralds sent	-	b
	292-297	N	8/26	Sacrifice and prayer	O	b
Il. 4	1-6	N	7/25	The gods in council drinking	-	b
	31-49	S	27/84	Zeus gives in to Hera, Troy shall fall	O	
	198/203	N	7/27	A herald goes to Machaon	-	b
	223-233	N	12/48	Agamemnon marshalls the troops	-	b
	250-256	N	8/27	Agam. goes through the troops	-	b
	292-302	N	16/51	Nestor marshals the troops	-	b
	326-337	N	16/50	Agam. goes to various men	-	b
	404-410	S	11/34	Sthenelos chides Agamemnon	O	
Il. 9	1-8	N	10/35	Panic grips Greeks; wind simile	O	x
	182-196	N	22/72	Od. and Ajax walk to Achilles' tent	-	b

Table 6. (continued)

Book	Lines	Mode	Coinc.	Sense of the Passages	Context	Type
Od. 4	113-137	N/N	40/121	Telemachus cries, Men. wonders. Helen enters; long description	(O)/-	(x)/b
	219-234	N	25/77	Helen's opiate described	-	b
	296-305	N	12/48	Preparations for retiring	-	a
	624-631	N	10/34	The suitors at play	-	b
	842-847	N	6/28	The suitors' plot; island described	-	a
Od. 7	153-158	N/N	7/26	Od. sits down; an old man speaks	-	b
	334-341	N	9/32	Preparations for retiring	-	a
*Od.*22	116-125	N	14/43	Od. shoots arrows, takes spears	O	x
	178-194	N	25/83	Mel. gets helmet; is caught, bound	O	x
	236-247	N/N	18/59	Athena sits watching; Agelaos rallies suitors	-	b
	465-473	N	13/45	The faithless maids are hanged, gaily	O	x
*Od.*23	174-180	S	11/33	Confused Pen. to Od. and Eurycleia	O	
	288-296	N	9/37	Bed preparations; Od. and Pen. retire	-	a

speech is hardly dramatic, especially when compared with Odysseus' vigorous retort (350 sq.). The first three-fourths of book 4 are, in general, not very vivid, hence the unusually large number of heterodyne passages. It is the lull before the storm of battle.

The situation at *Od.* 23. 174-180 is extraordinary. Penelope is confused, to judge from her syntax. Furthermore, the longer portion of her speech - and the most heterodyne - is simply a request that Eurycleia prepare the bed. This is a crucial point in the recognition, but Penelope must above all appear natural and calm when giving the order. Homer therefore used heterodyne texture.

Of the twenty-six narrative passages, seven describe preparatory activity of various kinds (type a); fifteen narrate activities which are not vigorous or dramatic (type b); four are apparently exceptional (type x). Despite these few exceptions the nature of the heterodyne passage is clear. Homodyne tends to be used for the dramatic, heterodyne for the undramatic.

In order to extend the coverage of this investigation into texture, portions of the works of Apollonius Rhodius, Quintus Smyrnaeus, and Nonnus were analyzed. This was originally done using the less precise method. Since use of the more precise method produces minimal changes in the results for Apollonius, I have not applied the more precise method to Quintus and Nonnus; for Apollonius I give the corrected results in parentheses (see Table 7). It is evident that, in later epic as in the *Iliad* and *Odyssey*, the percentage of homodyne lines is higher in speech than in narrative. The differential is, in fact, greater in later dactylic hexameter verse than in Homer: percentage of homodyne in speech in Homer - 29.1 % (29.2 %); in later epic - 34.2 %: percentage of homodyne in narrative in Homer - 25.8 % (25.2 %); in later epic - 25.5%. The differential has increased from 3.3 % (4.2 %) to 8.7 %.

Little need be said of the raw statistics beyond the salient fact that they support the conclusions reached pre-

Table 7. Homodynes in three different authors

Apollonius Rhodius *Argonautica* 3			
Speech lines:	552	Narrative lines:	855
Homodyne:	184 (185)	Homodyne:	226 (224)
Percentage:	33.3% (33.5%)	Percentage:	26.4% (26.2%)
Quintus Smyrnaeus 1, and 2 speech only			
Speech lines:	380	Narrative lines:	641
Homodyne:	117	Homodyne:	156
Percentage:	30.8%	Percentage:	24.3%
Nonnus *Dionysiaca* 1. 363-2 end			
Speech lines:	395	Narrative lines:	490
Homodyne:	153	Homodyne:	125
Percentage:	38.7%	Percentage:	25.5%
TOTAL			
Speech lines:	1327	Narrative lines:	1986
Homodyne:	454	Homodyne:	507
Percentage:	34.2%	Percentage:	25.5%

viously, that the more vivid mode of expression is the more homodyne. But *Argonautica* 3 warrants closer examination. As one scans the text observing the texture of each line, one is immediately aware that the distribution of homodyne lines is very uneven. When one returns to consider the context more carefully, the reasons for the changes in texture are readily apparent. It is not simply that speach is more homodyne than narrative, although that is generally true and significant. An analysis will show the expressional qualities of texture in *Argonautica* 3. (Where the more precise method changes the results, the new results will be given in parentheses.)

We begin with the speeches of Hera and Cypris (84-105 2/22). Hera speaks carefully or circumspectly (ἐπιφραδέως 83). She eschews force (84), and asks Cypris to give Eros his assignment softly (85). Note also εὐμενέουσα (87). Cypris' tale of Eros' naughty ways is not really very serious - Hera and Athena don't take the matter very seriously (100). Homodyne texture would be inappropriate for Hera's discreet diplomacy and for Cypris' complaints against her bothersome brat.

Much the same is true of Cypris' cajoling appeal to little Eros (129-144 3/16), a combination of mock reprimand and gentle bribery, much of it a description of the pretty toy. The few homodyne lines are among the most important and dramatic: 130 "You cheated him when he didn't know any better", 137 "It is made of gold hoops", 143 "shoot an arrow and make her fall in love with Jason." Notice in passing that 196-209 (1/14)(0/14) is simply a description of Jason and some companions on the way to town. The passage is of the sort that was common among the heterodyne passages in Homer.

The texture is heavily homodyne from 795-810 (12/16). In the first part of her soliloquy Medea has debated which is the lesser of the two evils. Now she imagines how the Colchian women would fault her if she went with Jason. She goes for the poison in tears, but is stopped by the fear of hateful Hades. Here is clearly a point of intense drama.

A lightly homodyne descriptive passage follows (825-866 7/42). We are told of Argus' departure and Medea's toilet. She calls her maids and prepares to drive to the Temple of Hecate. The magic ointment is described at length. The four lines which describe her washing seven times and calling on Brimo seven times are all completely heterodyne (860-63). Whether this is coincidence, or whether heterodyne texture was felt to be appropriate for such unnatural activities is uncertain. But when the earth quakes as the root is cut, it quakes in homodyne, the first homodyne after 15 non-homodyne lines. (864)

The first meeting between Jason and Medea is described in 948-974 (3/27). One might have anticipated a more homodyne texture, but a closer reading of the text shows what Apollonius was doing. I have previously made use of the "lull before

the storm" explanation, aware that the reader might dismiss it as "special pleading". Here is a clear example. The dominant image of the scene is of two tall trees standing rooted side by side in the still mountain air. So Medea was rooted to the spot, and she and Jason stood without speaking (967-70). When moved by the wind the tall trees make sounds without ceasing; even so, as Apollonius explicitly says, Medea and Jason *were about to* speak, stirred by the breath of Love (970-72). The texture expresses the essence of the scene.

Jason's speech follows (975-1007 7/33) (8/33). The texture is somewhat more homodyne, but still below the norm for speech in Apollonius. Jason is trying to put Medea at ease. He says that Medea need not be afraid, that they are friends. The speech is designed to reassure Medea and win her over. A heavily homodyne texture would be inappropriate to the mood of calm reassurance. The stronger lines are homodyne: 975 "Why are you so afraid of me, maiden?", 985 "I plead with you by Hecate", 989 "(without) you I shall not prevail in this terrible task", 992 "securing (you a) name and good reputation", 997 "Long ago (Ariadne) saved Theseus", 1005 "so you will have favor of the gods if you save".

From now on the speeches of Jason and Medea are more homodyne: esp. 1026-62 16/37 (15/37); 1069-76 5/8; and 1079-1101 11/23 (except for the description of Achaea which includes eight consecutive non-homodyne lines 1087-94).

Finally, there is a long descriptive passage which is, as usual, lightly homodyne (1176-1224 7/49). It begins with the past history of the dragon's teeth. Then follows a description of nightfall, the preparations for the sacrifice and the sacrifice, and the dawn. One might have thought that the arrival of Hecate Brimo and the chilling supernatural phenomena that accompany her appearance would have been expressed in more homodyne texture. Not only is the texture non-homodyne in general, but there is a series of three wholly heterodyne lines as before (1213-15, cf. 860-63). Wholly heterodyne sequences are very unusual (these two are, for example, the only wholly heterodyne sequences of more than two lines in *Argonautica* 3), but the evidence will probably not support the conclusion that heterodyne is the texture for the supernatural. In any case, the only homodyne line after a long series of non-homodyne lines is that which describes the meadow trembling under foot (1218), precisely as in the parallel passage a long series of non-homodyne lines was broken by one homodyne line which described the earth quaking. (864)

Before presenting the statistical evidence I stated that poetic rhythm is a matter of expectation, and that heaviness and stress are not basically different. We must now turn our attention to such matters as quantity (lightness or heaviness), stress, and the pitch accent, and their role in verse. Allow me to begin with two basic postulates offered by ALLEN (1964: 3). "(a) The underlying material of verse is language. Therefore we should not attribute to verse phenomena which are not present ... in the language upon which it is founded. (b) There is no special correlation between the chronological and typological classification of languages. Therefore, unless there is clear evidence to prove it, we should not attribute to a 'dead' language or its verse phenomena which are not observed in any living language or verse."

SEALEY (1963) has recently argued that, in Greek accentuation, moderately greater stress usually accompanied higher pitch. "In Lithuanian accentuation depends primarily on pitch, but a rise in pitch is usually accompanied by a moderate increase in stress (and a fall in pitch by a moderate decrease in stress)." The argument for applying this to Greek rests on a feature of Greek accentuation which is identified with a dialectal peculiarity of Lithuanian accentuation.

SEALEY is not alone. In a review of SZEMERÉNYI (1964) PULGRAM (1967) accepts SZEMERÉNYI's conclusions regarding pitch and stress: "Szemerényi is quite right in pointing out that one should not describe Greek as a pure pitch language and Latin as a pure stress language, since recent phonetic and acoustic investigations and experiments have shown that *such prosodic conditions do not exist*" (emphasis mine).

A new element, meter, may now be introduced into the discussion. In his discussion of the relationship between accent and meter ALLEN (1964) noted that, although Greek and Latin metrics were very similar in many respects, Greek did not show the relationship between verse-rhythm and word-accent which obtained in Latin.[9] That forced him to proceed backwards chronologically (p. 4). But there was a much greater problem. On pages 5 ff. ALLEN undertook a rather complicated argument: In the main meters common to Latin and Greek (dactylic and iambic) there is, in Latin, a beat which falls on the 'strong position'. Rules for the structure of the syllable in this strong position are the same as for the structure of accented syllables. Therefore the nature of ictus and accent must have been similar. And, since the rules for placing the ictus in Greek verse are based on the same principle (heavy quantity), the ictus must have been manifested by stress (or, to put it differently and in more general terms, heavy syllables were liable to receive stress).

ALLEN was willing to go that far in arguing from one language to the other, but he had to stop short of the final correlation and admit the one flaw: "In Greek, since the linguistic word-accent is tonal and the verse-ictus is dynamic, it is irrelevant to speak either of concord or of conflict between ictus and accent" (p. 7). But SEALEY has provided evidence that the Greek accent is partly dynamic, a point which follows logically from ALLEN's own postulates (and is accepted by A. SCHMITT, DEBRUNNER, SZEMERÉNYI and PULGRAM): and the present study supplies the missing evidence of relationship between verse-rhythm and word-accent. So it is possible to close the ring of evidence and say that there is no reason to doubt that stress, quantity, and accent are related in Greek.[10]

These various phenomena had been brought together previously. Victor ERLICH explains the Formalist TOMASHEVSKIJ's definition of rhythm ("the totality of actually perceptible phonic phenomena") so: "This obviously included both the 'quantitative' or 'relational' factors usually discussed under the heading of metrics, such as stress, pitch, and length, and the 'qualitative' or 'inherent' ones - alliteration or vowel harmony" (1965: 215, 219). He also gives JAKOBSON's views: "In general, Jakobson argued ..., language has at its disposal ... three potential 'bases of rhythm' - the stress or 'dynamic accent', the pitch or 'musical accent', and quantity. Actually, each individual language, at a given period of its evolution, favors one of the above elements as the organizing principle of its versification." He implies that, though one might predominate, the others may be present.

Regarding the conclusions reached so far two points should be made to forestall criticism or misunderstanding. First, I have no intention of challenging quantity as the (prime) determinant of Greek meter. But the degree of coincidence or non-coincidence changes the character of the line. Second, the orthodox view that there is no relationship between stress (or regularly-recurring heavy syllables) and accent in Greek - and that is precisely what is being challenged - is based on the assumption that such a relationship must appear, if it exists at all, as overwhelming coincidence. But the Greek poets were in no respect so ham-handed (nor was Vergil). One could as well assume that if strong sense-breaks were significant in stichic verse, then every dactylic hexameter must end with a strong break. In fact enjambment was sought to avoid such deadly monotony. More primitive poetry does not show enjambment, and if Homer were less sophisticated, we might well find a much higher percentage of coincidence between regularly-recurring heavy syllables and high pitch accent.

To summarize: W.F. Jackson KNIGHT has shown that there is some relationship between meter and accent in Latin hexameter verse. W.S. ALLEN argued that in both Greek and Latin "the rules governing the placing of the ictus ... are based on ... the incidence of heavy quantity ... If, therefore, the ictus of classical Latin verse was manifested by stress, it is reasonably certain that the same applied to Greek verse - i.e. that here also the rules of quantity are rules of stressability (ALLEN 1964: 6). He felt that he dare go no further in his correlation because he considered the Greek accent primarily tonal unlike the Latin dynamic accent, and because the evidence for correlation between verse ictus and accent in Latin was lacking for Greek hexameter poetry. But SEALEY's article (and the views of a number of linguists) require that the absolute distinction between the Greek and Latin accents be given up; and the present study supplies the missing evidence of correlation between regularly-recurring heavy syllables and accent. Therefore ALLEN's correlation can be extended, we must give up our neat, water-tight categories of stress, quantity, dynamic and pitch-accent, and we may conclude that in both Greek and Latin poetry, as in the poetry of other languages, the linguistic phenomena stress, quantity, and accent are interrelated.

NOTES

*I wish to thank Professor Wm. F. WYATT, Jr. for advice and helpful criticism. Professor John COLLINS of the Department of Mathematics and Statistics (University of Calgary), provided essential assistance and calculated the revised chi squares for samples BA and CA. Professors W. S. ALLEN, W. B. STANFORD, and J. B. McDIARMID deserve thanks for their interest and encouragement. The last refinement of this paper must be credited to Dr. R. GROTJAHN.

1 WILKINSON (1966: 120-134, 221-236); cf. WHATMOUGH (1956: 10): "Moreover it is clear that no pattern - Sanskrit, Greek and Latin included - is wholly quantitative, but that other factors of accent, pitch and word structure enter into any verse pattern."

2 *Rhet.* 1403 b 24-9 and *Poet.* 1456 b 30-4 as quoted and translated by STURTEVANT (1940: 99). Some scholars maintain that

Aristotle is referring to sound qualities, e.g. shrillness, in the first passage. It is difficult to imagine how Aristotle, with his cordial horror of ambiguity, could have described tone quality in precisely those terms which were used for the accents.

3 Compare Victor ERLICH's (1965: 13) definition of rhythm: "... the rhythm of a poem hinges not so much on the actual distribution of the rhythmical accents as on our anticipation of their recurrence at certain intervals. To JAKOBSON, as well as to WELLEK and WARREN, 'the time of verse language is the time of expectation', (p. 171)." Compare WILKINSON (1966: 95 f. and 145).

4 The effect of the long syllable is described by Demetrius, *On Style*, 39: "φύσει γὰρ μεγαλεῖον ἡ μακρά, καὶ προλεγομένη τε πλήσσει εὐθύς. ... For the long syllable is grand by its very nature, and striking when used straightway at the beginning." ALLEN (1964: 10) says: "... there is no reason to believe that the phenomena called 'stress' and 'quantity' are substantially different."

5 The reader may refer to SCHMIEL (1968) for a fuller discussion of various points and for line-by-line textural analysis of the twelve books in schematic form.

6 The refined method was used to check the original results in various ways: 1) sample BA was taken again with the following result; the original figure of 1088 homodyne lines in speech is corrected to 1089, and the original figure of 766 homodyne lines in narrative is corrected to 748. Thus the differential has increased, favourably for my thesis. 2) Sample CA was taken again, with a correction favourable to my thesis. 3) The sample from the *Argonautica* was retaken; the variation was again slight, but marginally favourable to my thesis.

7 WILKINSON (1966: 120-132 and Appendix I) discusses the problem in detail. WELLEK/WARREN (1956: 164): "Boris Tomashevsky ... has shown by statistical methods that the beginnings and ends of sentences tend towards greater rhythmical regularity than do the centers." The same appears to be true of the dactylic hexameter line.

8 There is no convenient way to figure the hypergeometric probability distribution of longer passages (because the number of syllables in even a seven-line passage will go beyond the maximum value of N calculated by LIEBERMAN/OWEN, that is 100). It has been shown that the more precise method produces very small changes in the results. One can assume that, the larger the sample, the less significant these changes will be (cf. in this volume R. GROTJAHN, A Statistical Model for the Analysis of the Coincidence of Ictus and Accent, Part 3.1). I have, therefore, used the results obtained by the less precise (but more convenient) method for studying heterodyne passages.

9 The relationship is hardly more obvious in Vergil than in Homer. The high degree of coincidence between ictus and accent in Roman Comedy might have been necessitated by the large number and variety of resolutions, or it could be considered an example of Roman Comedy's unsophistication.

10 It must be stated that Professor ALLEN has come to very different conclusions. He has argued that Greek had a stress (which fell primarily on the last heavy syllable of polysyllabic words) as well as a pitch accent; that this stress was closely correlated with verse ictus (ALLEN 1966); and that there is a correlation between this stress and svarita in Greek (ALLEN 1967). See now ALLEN 1973. A discussion of these provocative and closely-argued theories would extend this paper to impossible length. The reader may refer to NEWTON (1969) and to BAILEY's review of ALLEN's *Vox Graeca* (1969) for arguments against this particular theory of ALLEN's.

A P P E N D I X

A SELECTION OF PERTINENT EXAMPLES FROM THE LIEBERMAN-OWEN TABLES OF THE HYPERGEOMETRIC PROBABILITY DISTRIBUTION

A) *To determine homodyne lines*: the upper limit for homodyne is .300.

Syl.	Accents	Coinc.	P(x)
14	3	1	.846*
14	4	1	.930
14	5	1	.972
15	3	1	.815
15	4	1	.908
15	5	1	.958
15	6	1	.983
16	3	1	.786
16	4	1	.885
16	5	1	.942
16	6	1	.974
17	3	1	.757
17	4	1	.861
17	5	1	.925
17	6	1	.963
14	3	2	.385
14	4	2	.594
14	5	2	.762
14	6	2	.879
15	3	2	.341
15	4	2	.538
15	5	2	.706
15	6	2	.832
15	7	2	.916
16	3	2	.304
16	4	2	.489
16	5	2	.653
16	6	2	.785
16	7	2	.879
17	3	2	.272
17	4	2	.445
17	5	2	.605
17	6	2	.739
17	7	2	.840

Syl.	Accents	Coinc.	P(x)
14	4	3	.175
14	5	3	.343
14	6	3	.529
14	7	3	.704
15	4	3	.143
15	5	3	.287
15	6	3	.455
15	7	3	.622
16	4	3	.118
16	5	3	.242
16	6	3	.392
16	7	3	.549
16	8	3	.696
17	4	3	.099
17	5	3	.205
17	6	3	.339
17	7	3	.484
17	8	3	.627
14	5	4	.063
14	6	4	.156
14	7	4	.296
15	5	4	.047
15	6	4	.119
15	7	4	.231
15	8	4	.378
16	5	4	.036
16	6	4	.092
16	7	4	.182
16	8	4	.304
16	9	4	.451
17	5	4	.028
17	6	4	.072
17	7	4	.145
17	8	4	.247
17	9	4	.373

Syl.	Accents	Coinc.	P(x)
14	6	5	.016
14	7	5	.051
14	8	5	.121
15	6	5	.011
15	7	5	.035
15	8	5	.084
16	6	5	.008
16	7	5	.024
16	8	5	.059
16	9	5	.121
17	6	5	.005
17	7	5	.018
17	8	5	.043
17	9	5	.088

Syl.	Accents	Coinc.	P(x)
14	7	6	.002
14	8	6	.009
15	7	6	.001
15	8	6	.006
15	9	6	.017
16	7	6	.0009
16	8	6	.003
16	9	6	.010
16	10	6	.026
17	7	6	.0006
17	8	6	.002
17	9	6	.007
17	10	6	.017

*(The values from the Lieberman and Owen tables are here presented in a form most convenient for a classicist who is looking at a text of Homer. The first example will appear in Lieberman and Owen under N=14, n=11, k=6, x=5).

The refined method of determining homodyne texture will produce results different from those produced by the less precise method in these cases:

1) lines which show four coincidences per seven acute and circumflex accents will now be homodyne when the line has between 14 and 17 syllables.
2) lines which show three coincidences per five acute and circumflex accents will be homodyne (as before) when the line has between 15 and 17 syllables; non-homodyne (now) when the line has 14 syllables or less.
3) lines which show two coincidences per three acute and circumflex accents will be homodyne (as above) when the line has 17 syllables; non-homodyne (now) when the line has fewer than 17 syllables.

The P(x) value for lines which show no coincidences will always be 1.000.

B) *To determine heterodyne lines:* the upper limit for heterodyne is .700. A selective sample will illustrate how the Lieberman-Owen tables are used to determine heterodyne lines. Lines which show 4 coincidences and more will never be heterodyne.

Syl.	Accents	Coinc.	P(x)
14	6	3	.844*
14	7	3	.704
14	8	3	.529
15	7	3	.769
15	8	3	.622
16	7	3	.818
16	8	3	.696
17	7	3	.855
17	8	3	.753
14	5	2	.657
14	6	2	.471
15	5	2	.713
15	6	2	.545
16	5	2	.758
16	6	2	.608
17	5	2	.795
17	6	2	.661

Syl.	Accents	Coinc.	P(x)
14	3	1	.615
14	4	1	.406
14	5	1	.238
15	3	1	.659
15	4	1	.462
15	5	1	.294
16	2	1	.875
16	3	1	.696
16	4	1	.511
16	5	1	.346
16	6	1	.215
17	2	1	.890
17	3	1	.728
17	4	1	.555
17	5	1	.395
17	6	1	.261

*(The values from the Lieberman and Owen tables are here presented in a form most convenient for a classicist who is looking at a text of Homer. The first example will appear in Lieberman and Owen under N=14, n=6, k=6, x=3.)

The only cases where the results will differ from those obtained by the former, less precise method, are the following:

1) lines which show one coincidence per three acute and circumflex accents will be heterodyne (as before), except when the line has 17 syllables.
2) lines which show two coincidences per five acute and circumflex accents will now be heterodyne if the line has 14 syllables or less.
3) lines which show three coincidences per eight acute and circumflex accents will be heterodyne (as before) except when the line has 17 syllables.

References

ALLEN, W.S., On Quantity and Quantitative Verse. In: ABERCROMBIE, D. et al. (Eds.), In Honour of Daniel Jones. London 1964

ALLEN, W.S., Prosody and Prosodies in Greek. Transactions of the Philological Society. Oxford 1966, 107-148

ALLEN, W.S., Correlations of Tone and Stress in Ancient Greek. In: To Honour Roman Jakobson. The Hague 1967, 46-62

ALLEN, W.S., Accent and Rhythm. Cambridge 1973

BAILEY, C.-J.N., Review of W.S. Allen, Vox Graeca. Language 45, 1969, 644-646

ERLICH, V., Russian Formalism: History-doctrine. The Hague 1965

GETTY, R.J., Classical Latin Metre and Prosody 1935-1962. Lustrum 8, 1963, 103-160

KNIGHT, W.F.J., Accentual Symmetry in Vergil. Oxford 1939

KNIGHT, W.F.J., Roman Vergil. London 1944

LIEBERMAN, G.J. / OWEN, D.B., Tables of the Hypergeometric Probability Distribution. Stanford 1961

MAAS, P., Greek Metre. Transl. H. LLOYD-JONES. Oxford 1962

NEWTON, B.E., Metre and Stress in Greek. Phoenix 23, 1969, 359-371

PULGRAM, E., Review of O.J.L. Szemerényi, Syncope in Greek and Indo-European, and the Nature of Indo-European Accent. American Journal of Philology 68, 1967, 218-223

SCHMIEL, R., Rhythm and Accent in Homer. Ann Arbor, University Microfilms 1968

SEALEY, R., Stress as a Factor in Classical Greek Accentuation. Greece and Rome 10, 1963, 11-25

STURTEVANT, E.H., The Pronunciation of Greek and Latin. Philadelphia 1940

SZEMERÉNYI, O.J.L., Syncope in Greek and Indo-European, and the Nature of Indo-European Accent. Naples, Istituto Universitario Orientale di Napoli 1964 (Quaderni della Sezione Linguistica degli Annali, No. III)
WELLEK, R. / WARREN, A., Theory of Literature. New York 1956
WHATMOUGH, J., Poetic, Scientific and Other Forms of Discourse. Berkeley 1956
WILKINSON, L.P., Golden Latin Artistry. Cambridge 1966

A Statistical Model for the Analysis of the Coincidence of Ictus and Accent

Rüdiger Grotjahn, Bochum

0. Introduction

In the present article we will first show how by means of some relatively simple combinatorial deliberations one can arrive at the hypergeometric probability distribution (henceforward HPD) as a model for the coincidence of ictus and accent in hexameter verse and in other kinds of verse as well. After describing some characteristics of this distribution which are very useful for practical applications, we will discuss an alternative derivation of the HPD. We will then present a table permitting an objective classification of hexameter lines into homodyne, heterodyne and neutral. After that we will deal with some approximations to the HPD such as the binomial distribution and the normal distribution and their application in the analysis of longer passages of verse. Finally we will present a coefficient which allows us to measure the degree of association between the variables 'ictus' and 'accent'.

In order to facilitate understanding, all statistical procedures are illustrated by at least one application. We hope that in this way the procedures discussed can be made accessible to as many researchers as possible.

1. The Hypergeometric Probability Distribution

1.1. Derivation and Some Characteristics

Although the HPD has a wide range of application in other disciplines it has not been used very often so far in language-oriented disciplines such as linguistics, poetics or versification. Particularly in the field of versification the HPD seems to have potentially very fruitful applications. One can for instance compute the exact probability of the coincidence of ictus and accent and thus arrive at an objective classification of verse lines according to whether they are homodyne, heterodyne or neutral (see below p. 46 and SCHMIEL in this volume). This classification procedure can be extended to other versification systems such as Latin elegiac verse and also to whole texts.

The HPD can be derived as follows: if there are N syllables in a line, out of which n are accented and k are heavy, the total number of possibilities of choosing n accented syllables among the N syllables is

$$\binom{N}{n}. \tag{1}$$

Since the number of possibilities of getting x heavy syllables which are also accented corresponds to the number of possibilities of choosing x accented syllables among the k heavy syllables and n - x unaccented syllables among the N - k light syllables we have

$$\binom{k}{x}\binom{N-k}{n-x} \tag{2}$$

possibilities of getting a coincidence between heavy syllables and accents.

The probability that exactly x heavy syllables and accents coincide is now given as the number of favourable events divided by the number of possible events. Since the number of favourable events is given by (2) and the number of possible events by (1) we obtain

$$P(X=x) = f(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \tag{3}$$

with $\max(0,n+k-N) \le x \le \min(n,k)$,

which is known as the hypergeometric probability distribution.

The corresponding cumulative distribution is

$$P(X \le x_c) = H(x) = \sum_{x=u}^{x_c} \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \tag{4}$$

with $u = \max(0,n+k-N)$ and $x_c \le \min(n,k)$.

Example 1. Let us illustrate why $\max(0,n+k-N) \le x \le \min(n,k)$ by means of an example. Suppose we had observed a hexameter line with

N = 16 syllables
n = 12 accented syllables
k = 6 heavy syllables.

In this line the lower limit for the number of coincidences of accented and heavy syllables must be $\max(0,12+6-16) = 2$ since there are only 4 unaccented syllables which can coincide with at most 4 heavy syllables. Consequently, 2 heavy syllables must necessarily coincide with two accented syllables. Since 6 heavy syllables can coincide with at most 6 accented syllables the upper limit for the number of coincidences must be $\min(6,12) = 6$.

For some applications it is important to know the mean and the variance of the HPD. These two characteristics can be most easily derived using the method of factorial moments (cf. MOOD/GRAYBILL/BOES 1974: 77 and 91f.). Since the derivations are somewhat tedious we only present the results. Setting $p = k/N$ and $q = 1-p$ we get for the mean

$$E(X) = \mu_1' = \frac{nk}{N} = np \tag{5}$$

and for the variance

$$V(X) = \mu_2 = \frac{nk(N-n)(N-k)}{N^2(N-1)} = npq\,\frac{N-n}{N-1}. \tag{6}$$

The computation of the cumulative HPD is sometimes very cumbersome. Fortunately, there exist extensive tables by LIEBERMAN/OWEN (1961). The use of these tables involves, however, several identities. For $N \leq 25$, which is always the case when we consider only single hexameter lines, the only identity that has to be used in entering the tables is

$$\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = \frac{\binom{n}{x}\binom{N-n}{k-x}}{\binom{N}{k}} . \tag{7}$$

Because of this very important identity, the HPD is tabulated only for $k \leq n$. If $k > n$, one merely has to interchange n and k to read the probabilities directly from the tables.

Example 2. In *Iliad* 1, 237 we have (cf. SCHMIEL in this volume p. 9):

N = 15 syllables
n = 7 accented syllables
k = 6 heavy syllables
x = 4 accented and heavy syllables.

In order to obtain the probability of exactly 4 coincidences of heavy syllables and accents, we need only look under N = 15, n = 7, k = 6 and x = 4 in the LIEBERMAN/OWEN tables to obtain

$$P(X = 4) = \frac{\binom{6}{4}\binom{15-6}{7-4}}{\binom{15}{7}} = 0.195804.$$

Example 3. For *Iliad* 1, 238, SCHMIEL (ibid.) found the following values:

N = 17, n = 5, k = 6, x = 1.

We now want to look for the probability $P(X \leq 1;\ N=17,\ n=5,\ k=6)$. Since $k > n$, we have to interchange n and k (cf. formula 7). We therefore enter the table under

N = 17, n = 6, k = 5, x = 1

and find

$$P(X \leq 1) = \sum_{x=0}^{1} \frac{\binom{6}{x}\binom{17-6}{5-x}}{\binom{17}{5}} = \sum_{x=0}^{1} \frac{\binom{5}{x}\binom{17-5}{6-x}}{\binom{17}{6}} = 0.394635.$$

Example 4. In *Iliad* 1, 241 we have (cf. SCHMIEL, ibid.) N = 16, n = 6, k = 6. We want to compute $P(X \geq 4)$. Since in the LIEBERMAN/OWEN tables only $P(X \leq x)$ is tabulated, we use the following identity, which generally holds for discrete distributions:

$$P(X \geq x) = 1 - P(X \leq x-1). \tag{8}$$

We get

$$P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.908092 = 0.091908.$$

It is, however, possible to read $P(X \geq x)$ directly from the LIEBERMAN/OWEN tables as well. Since asking for the probability that out of n accents and k heavy syllables x or more accents and heavy syllables coincide is logically equivalent to asking for the probability that among N-n unaccented syllables and k heavy syllables k-x or less unaccented syllables are heavy, the following identity holds:

$$P(X \geq x;\ N,n,k) = P(X \leq k-x;\ N,N-n,k), \tag{9a}$$

which can also be written by interchanging k and n as

$$P(X \geq x;\ N,n,k) = P(X \leq n-x;\ N,N-k,n). \tag{9b}$$

Because of this identity

$$P(X \geq 4;\ 16,6,6) = P(X \leq 2;\ 16,10,6).$$

We therefore enter the table under

$$N = 16,\ n = 10,\ k = 6,\ x = 2$$

and get 0.091908 as before. It should be noted that the identity (9) is used by SCHMIEL (in this volume) in classifying lines as homodyne (e.g. p.8).

If tables of the HPD are not at hand, the calculation of the cumulative distribution can be facilitated by a recursion formula which is derived by first calculating

$$\frac{P(X=x+1)}{P(X=x)} = \frac{\binom{k}{x+1}\binom{N-k}{n-x-1}}{\binom{N}{n}} \div \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = \frac{(k-x)(n-x)}{(x+1)(N-k-n+x+1)}$$

and then solving for $P(X = x+1)$:

$$P(X=x+1) = \frac{(k-x)(n-x)}{(x+1)(N-k-n+x+1)} \cdot P(X=x). \tag{10}$$

<u>Example 5</u>. We calculate $P(X \geq 4;\ N=16,\ n=6,\ k=6)$ from Example 4 by using the recursion formula (10). We first compute

$$P(X=0) = \frac{\binom{6}{0}\binom{16-6}{6-0}}{\binom{16}{6}} = 0.026224,$$

then

$$P(X=1) = \frac{(6-0)(6-0)}{1(16-6-6+1)} \cdot 0.026224 = 0.188811$$

$$P(X=2) = \frac{(6-1)(6-1)}{2(16-6-6+2)} \cdot 0.188811 = 0.393357$$

$$P(X=3) = \frac{(6-2)(6-2)}{3(16-6-6+3)} \cdot 0.393357 = 0.299700.$$

Using the identity (8) we finally obtain

$$P(X \geq 4) = 1 - P(X \leq 3) = 1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3)] =$$
$$= 0.091908.$$

1.2. An Alternative Approach

If we represent the relation between ictus and accent in e.g. *Iliad* 1, 237 (cf. Example 2) in the form of a 2x2 contingency table,we get Table 1.

Table 1. Relation between Ictus and Accent in *Iliad* 1, 237

Syllables	heavy	light	
accented	4	3	7
unaccented	2	6	8
	6	9	15

Obviously the classification shown in Table 1 is only a special case of a cross classification of N elements with respect to two properties that any element may or may not possess. The result of such a cross classification with regard to two dichotomous properties can be represented in a general form as in Table 2.

Table 2. General Form of a 2x2 Table

	B	$\bar{B}$	
A	a	b	$n = a+b$
$\bar{A}$	c	d	$r = c+d$
	$k=a+c$	$s=b+d$	$N = a+b+c+d$

It is easily seen that if we consider the marginal totals in Table 2 as fixed[1] the probability of getting a elements which have both properties A and B is again given by the HPD (3) which now can be written as

$$f(a) = \frac{\binom{a+c}{a}\binom{N-a-c}{b}}{\binom{N}{a+b}} =$$

$$= \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{N!}\,\frac{1}{a!b!c!d!}. \qquad (11)$$

Thus the HPD is not only a model for the analysis of the coincidence of ictus and accent but can be considered as a *general* model for the analysis of the relation between two dichotomous properties.

Since in Table 2 the marginal totals are fixed, knowing the value for one cell implies that we also know the values for the remaining cells. Therefore, in the case of a 2x2 table like Table 2 the following identities hold (cf. the identities (7) and (9)):

$$\begin{aligned} f(a;\,N,n,k) &= f(b;\,N,n,N-k) \\ &= f(c;\,N,k,N-n) \\ &= f(d;\,N,N-n,N-k) \end{aligned} \qquad (12a)$$

Since one may always interchange rows and columns in a table such as Table 2, we may also interchange n and k. Hence the following identities hold as well:

$$\begin{aligned} f(a;\,N,k,n) &= f(b;\,N,k,N-n) \\ &= f(c;\,N,n,N-k) \\ &= f(d;\,N,N-k,N-n) \end{aligned} \qquad (12b)$$

Example 6. The application of the above formulas is now illustrated by means of the data of Table 1. Using (11) and (12a) we get

$$f(4;\,15,7,6) = f(3;\,15,7,9) = f(2;\,15,8,6) = f(6;\,15,8,9) =$$

$$= \frac{7!8!6!9!}{15!}\,\frac{1}{4!3!2!6!} = 0.195804,$$

which is the same result as obtained in Example 2.

If we want to test two dichotomous properties such as ictus and accent for independence, we are not interested in the probability that exactly *a* elements have both properties but in the probability of getting a distribution which is equal to or more extreme than the observed one. If we want to carry out a two-tailed test the probabilities we have to compute are in both tails of the HPD. In this case we test the null hypothesis (H_o) of no association between ictus and accent against the alternative hypothesis (H_A) that ictus and accent are associated. In the case of a one-tailed test H_A is split into a hypothesis of positive association (H_{A1}) and one of negative association (H_{A2}). In the case of H_{A1} the probabilities we have to compute are in the right hand tail of the HPD and in the case of H_{A2} they are in the left hand tail.

If we designate the proportion of coincidences expected under H_o by p_o the above hypotheses can be stated as follows:

H_o: $p = p_o$

H_A: $p \neq p_o$ (two-tailed test)

H_{A1}: $p > p_o$ (one-tailed test; positive association)

H_{A2}: $p < p_o$ (one-tailed test; negative association)

In our model p_o is given by the proportion $k/N = (a+c)/N$ (cf. formula 5), and p is to be estimated by the proportion $\hat{p} = x/n = a/(a+b)$. Under the null hypothesis of no association between ictus and accent we would expect the proportions k/N and x/n to differ only by chance. If there is a positive association we would expect x/n to be significantly larger, and if there is a negative association to be significantly smaller than k/N.

Let us once again consider Table 1. In this table $\hat{p} = 4/7 = 0.57$ and $p_o = 6/15 = 0.40$. Thus the observed proportion of coincidences is larger than we would have expected under H_o. If we want to carry out a one-tailed test we can proceed as follows:

Let the significance level to be chosen be $\alpha = 0.05$. Under the alternative hypothesis H_{A1} of positive association we have the following two tables, which both show distributions more extreme than that in Table 1.

5	2	7
1	7	8
6	9	15

6	1	7
0	8	8
6	9	15

With the help of (11) we now calculate

$$P(a \geq 4) = \frac{7!8!6!9!}{15!}\left(\frac{1}{4!3!2!6!} + \frac{1}{5!2!1!7!} + \frac{1}{6!1!0!8!}\right) = 0.230769,$$

which is, according to (12a), the same result as if we had calculated $P(c \leq 2)$ or $P(b \leq 3)$ or $P(d \geq 6)$. Since the value obtained is larger than α, we have to accept the null hypothesis of no association of ictus and accent in Table 1.

Let us now assume that we had observed a table with the same marginals as above but with $a = 6$. If we test the hypothesis H_{A1} of a positive association in this table on the $\alpha = 0.01$ level, we obtain $P(a \geq 6) = 0.001399 < 0.01$ and thus

have to accept H_{A1}. As the example shows, however, the actual significance level is only 0.0014 and thus far below the nominal level of 0.01. Because of the discreteness of the HPD, the nominal significance level can usually not be attained and hence the use of the HPD leads in many cases to a rather conservative decision, in particular if N is so small, as is the case in the analysis of single hexameter lines.[2]

As the above example already indicates, the computation of the cumulative HPD may be rather tedious even in the case of a 2x2 table, although the first term on the right hand side of (11) is a constant and thus has to be calculated only once. Therefore it is often advisable, in particular if one wants to analyse not only single hexameter lines but longer passages of hexameter verse, not to proceed as shown above but to use the recursion formula (10) or a special recursion formula for 2x2 tables (see FELDMAN/KLINGER 1963; JOHNSON 1972). Furthermore there exist extensive tables by FINNEY et al. (1963) and BENNETT/HORST (1966) especially designed for the application of the HPD of 2x2 tables.

The test of a 2x2 table we have just discussed is known as the Fisher-Yates test because it was devised concurrently by FISHER (1935) and YATES (1934). It is also referred to as "Fisher's exact test" because in contradistinction to the well known χ^2 test, which we will discuss in Part 3.3 and Part 4 of this article, it is an exact and not an approximative procedure for analyzing 2x2 tables.

Although the Fisher-Yates test has been originally devised as a one-tailed test it can also be used as a two-tailed test. There are, however, some problems connected with this latter use. In carrying out a two-tailed test with the help of the normal distribution for instance, we usually proceed in the following way: since the normal distribution is symmetric and thus $P(Z \geq z) = P(Z \leq -z)$, we only need to compute $2P(Z \geq z)$ *or* $2P(Z \leq -z)$ in order to obtain the two-tailed probability $P\{(Z \geq z) \cup (Z \leq -z)\} = P(Z \geq z) + P(Z \leq -z)$. If $2P(Z \geq z) = 2P(Z \leq -z)$ is greater than α or, which is the same, if $P(Z \geq z) = P(Z \leq -z)$ is greater than $\alpha/2$, we accept H_o; if not, we accept H_A.

If we proceed in the same way in the case of the Fisher-Yates test, a considerable over- or underestimation of the rejection error α may be the result, because the HPD is only symmetric if n or k is equal to $\frac{1}{2}N$. Let us illustrate this fact by the data of Table 1. Table 3 shows the values of P(a; 15,7,6) for a = 0,...,6.

Table 3. Values of P(a; 15,7,6) for a = 0,...,1 (cf. Table 1)

a	P(a)
0	0.005594
1	0.078322
2	0.293706
3	0.391608
4	0.195804
5	0.033566
6	0.001399

One realizes at once that there is a considerable difference between the probabilities in the lower and in the upper tail of the distribution P(a). Let us consider once again the case a = 6. If we carry out a two-tailed test with $\alpha = 0.01$ calculating only $P(a\geq 6)$ as proposed by, e.g., FINNEY (1948: 156) or SIEGEL (1956: 100), this would result in accepting H_A since $P(a\geq 6) = 0.00139 < \alpha/2 = 0.005$. If we had however calculated $P(a\leq 0)$ as well, we would have obtained a value of 0.005594, which is greater than $\alpha/2$. Thus, if we follow the convention usually adopted in the case of a two-tailed test that both the lower tail and the upper tail probabilities should be smaller or equal to $\alpha/2$ in accepting H_A, we now have to accept H_O. Hence, a decision based only on the calculation of $P(a\geq 6)$ would not be correct in this case.

As the example just discussed already indicates, the two-tailed test based on equal tail areas may in many cases be extremely conservative because of the discreteness of the HPD. The power of the test can, however, very often be increased by abandoning the convention of equal tail areas and proceeding in the following way as proposed by FREEMAN/HALTON (1951): First calculate P(a; N,n,k) for all possible values of a and fixed values of N, n and k. Then sum up all the probabilities which are less than or equal to P(a*) where a* designates the *a* which has been observed in the table under investigation. If this sum (which we may designate by P*) is less than or equal to α, accept H_A; otherwise accept H_O.[3]

Example 7. The procedure described is illustrated by the data in Table 1. Let us assume that a* = 5 and $\alpha = 0.05$. We calculate

$$P^* = P(a\geq 5) + P(a = 0) = 0.033566 + 0.001399 + 0.005594 = 0.040559$$

and accept H_A since $P^* < \alpha$.

If we had followed the convention of equal tail areas, however, we would have had to accept H_O since $P(a\geq 5) = 0.034965 > \alpha/2 = 0.025$. In the present case the Freeman-Halton procedure thus leads to an enlargement of the rejection region and is consequently more powerful than the two-tailed Fisher-Yates test based on equal tail areas.

If *a* can take a large number of values and not only six, as is the case in the analysis of single hexameter lines, the Freeman-Halton procedure may result in a considerable amount of computation. There are however several possibilities to reduce the number of calculations involved (see FREEMAN/HALTON 1951, JOHNSON 1972 and WALL 1972). Furthermore there exist tables developed by WALL and printed in LIENERT (1975: 442ff) which give the two-tailed probabilities as defined by FREEMAN/HALTON (1951) for N = 6(1)25 and $p \leq 0.20$ as well as for N = 25(5)60, $N-n \leq N/5$, $N-k \leq N/5$, and $p \leq 0.20$.

2. *Classifying Hexameter Verse with the Help of the Hypergeometric Probability Distribution*

2.1. We now want to show how to use the HPD in classifying hexameter verse according to the number of coincidences of accents and heavy or light syllables. To this end we first discuss the classification attempt by SCHMIEL (in this volume).

SCHMIEL proceeds as follows: he sets the upper limit for homodyne at P = 0.30 and that for heterodyne at P = 0.70 on the LIEBERMAN/OWEN tables. (Lines which show a tendency towards coincidence of ictus and accent are called homodyne and lines which do not show such a coincidence are called heterodyne.) Lines which are neither homodyne nor heterodyne are considered as neutral (see p. 8 and p. 11). In order to be able to read the probabilities for both homodyne and heterodyne directly from the LIEBERMAN/OWEN tables, in the case of homodyne he designates by n (as already mentioned) not the number of accented syllables but the number of unaccented syllables.

This procedure is open to criticism from several points of view:

(1) SCHMIEL does not give any argument as to why he accepts a greater probability of making a false decision in the case of heterodyne lines than in that of homodyne lines. In my opinion there is no reason why the probability of making a false decision should not be the same in both cases.

(2) Choosing 0.30 for homodyne and 0.70 for heterodyne results in a very low probability of classifying a line as neutral. This probability may in fact be equal to zero and can only be greater than zero because the HPD is a discrete distribution.

(3) An upper limit of P = 0.70 for heterodyne may lead to the paradoxical consequence that we have to classify a line as heterodyne although the observed number of coincidences is <u>not</u> smaller than the number expected under H_o. To illustrate this let us consider a line with N = 16 syllables, n = 8 accents and x = 3 coincidences. In this line E(X) = nk/N = 3 and thus the observed number of coincidences is equal to the number expected under H_o. Consequently, ictus and accent should be considered as completely unrelated. If we adopt, however, SCHMIEL's criterion, the line has to be classified as heterodyne because $P(X \leq 3;\ 16,8,6) = 0.695804$ and $P(X \geq 3;\ 16,8,6) = 0.304196$ (cf. also SCHMIEL's Table B on p. 30 in this volume).

(4) In classifying lines as homodyne or heterodyne we assume that the structure of a line is not only the product of chance but rather of a tendency toward structuring on the part of the poet. A probability of $P(X \leq x;\ N,n,k) = 0.70$ on the LIEBERMAN/OWEN tables means, however, that if we draw a very large number of *random* samples from a universe of N syllables out of which n are accented and N-n unaccented we obtain a coincidence of x or less heavy syllables and accents in about 70% of the samples, <u>this coincidence being merely due to chance</u>. An equivalent argument holds for the probability of 0.30 in the case of homodyne lines. SCHMIEL thus accepts a probability of making an error of up to 0.70 in classifying verses as heterodyne and of up to 0.30 in classifying verses as homodyne. Such a high probability of making a false decision does not seem acceptable.

Taking this criticism into account, we now propose the following general model for the classification of hexameter lines into homodyne, heterodyne and neutral:

If prior to the collection of the data we have no justified hypothesis in which direction the observed proportion $\hat{p} = x/n$ will differ from the proportion $p_o = k/N$ expected under H_o, we carry out a two-tailed test. If, as in Part 1.2 of this article, P* denotes the sum of the probabilities which are less than or equal to $P(X=x_c)$ where x_c designates the number of coincidences observed, the classification is as follows:

$$\begin{array}{lll} \text{heterodyne:} & P^* \leq \alpha & \text{and } \hat{p} < p_o \\ \text{homodyne:} & P^* \leq \alpha & \text{and } \hat{p} > p_o \\ \text{neutral:} & P^* > \alpha & \end{array} \qquad (13)$$

If we have a justified hypothesis concerning the direction in which $\hat{p}$ will differ form p_o, we carry out a one-tailed test. In this case we classify as follows:

$$\begin{array}{lll} \text{heterodyne:} & P(X \leq x_c) \leq \alpha & \text{and } \hat{p} < p_o \\ \text{homodyne:} & P(X \geq x_c) \leq \alpha & \text{and } \hat{p} > p_o \\ \text{neutral:} & P(X \leq x_c) > \alpha & \text{and } \hat{p} \leq p_o \\ \quad\text{or} & P(X \geq x_c) > \alpha & \text{and } \hat{p} \geq p_o \end{array} \qquad (14)$$

The point is now how to choose α. As we have seen, in the classification proposed by SCHMIEL, α (virtually) amounts to $0.70 + 0.30 = 1$, i.e. to 100%. In statistical research it is however a convention that α should not be greater than 0.10 and very often only $\alpha = 0.05$ is accepted. Nevertheless, since it finally depends on the objective of a study which level of significance is to be chosen, there is no general solution to this problem. Furthermore, since in the analysis of hexameter verse erroneous acceptance or rejection of the null hypothesis usually does not have such serious consequences as in, say, medical research, we suggest accepting significance levels which are even greater than 0.10. If we classify single hexameter lines, this is the only way to increase the power of our test, because N cannot be increased.

Example 8. Let $\alpha = 0.20$. Without taking the grave into account, we get the following two-tailed probabilities P* and one-tailed probabilities $P(X \leq x_c)$ or $P(X \geq x_c)$ for *Iliad* 237-241 (cf. SCHMIEL, in this volume p. 9):

237: $x_c=4$, N=15, n=7, P*=0.314685, $P(X \geq 4)=0.230769$

238: $x_c=1$, N=17, n=5, P*=0.600033, $P(X \leq 1)=0.394635$

239: $x_c=2$, N=17, n=5, P*=1.000000, $P(X \geq 2)=0.605365$

240: $x_c=3$, N=16, n=5, P*=0.299451, $P(X \geq 3)=0.241759$

241: $x_c=4$, N=16, n=6, P*=0.118133, $P(X \geq 4)=0.091909$

According to both criteria (13) and (14), line 241 must be considered as homodyne, and the lines 237-240 as neutral. The hypothesis that there is a tendency toward structuring in the use of accents and heavy syllables is thus accepted only in the case of line 241. According to the criterion proposed by SCHMIEL, however, the lines must be classified as follows:

heterodyne:	238
homodyne:	237, 240, 241
neutral:	239

We see that SCHMIEL's criterion results in a considerably different classification. If we choose $\alpha = 0.10$ instead of $\alpha =$ $= 0.20$ and if we have no justified hypothesis on the direction of the difference between $\hat{p}$ and p_o, then even line 241 has to be classified as neutral.

2.2. In order to facilitate the use of the HPD in the analysis of hexameter verse, I have prepared a table (see Table 4), which allows quick and simple classification of hexameter lines according to our criteria (13) and (14). In this table one finds for all combinations of N, n and k possible in single hexameter lines those probabilities $P(X=x) = f(x)$ which are less than or equal to 0.20. Since all combinations which are *theoretically* possible are given, at least some of the combinations taken into account will probably not occur in hexameter texts. To make the use of the table as easy as possible I have not eliminated entries which may be considered as redundant because of the identities given in Part 1.1 of this article. In order to obtain P*, $P(X \leq x_c)$ or $P(X \geq x_c)$ from Table 4 we only need to add the corresponding entries. Since Table 4 gives all values of f(x) which are less than or equal to 0.20 one can choose any significance level ≤ 0.20. In most cases one can even use significance levels far greater than 0.20.

Example 9. To illustrate the use of Table 4, we consider once again *Iliad* 241 (cf. Example 8). Since we had no hypothesis as to whether $\hat{p} > p_o$ or whether $\hat{p} < p_o$ prior to the collection of the data, we use the criterion (13). Let $\alpha = 0.20$.

Table 4. Hypergeometric Probabilities for $12 \leq N \leq 17$, $k = 6$ and $f(x) \leq 0.20$.

N	n	x	f(x)
12	3	0	.090909
	3	3	.090909
	4	0	.030303
	4	4	.030303
	5	0	.007576
	5	1	.113636
	5	4	.113636
	5	5	.007576
	6	0	.001082
	6	1	.038961
	6	5	.038961
	6	6	.001082
	7	1	.007576
	7	2	.113636
	7	5	.113636
	7	6	.007576
	8	2	.030303
	8	6	.030303
	9	3	.090909
	9	6	.090909
13	2	2	.192308
	3	0	.122378
	3	3	.069930
	4	0	.048951
	4	3	.195804
	4	4	.020979
	5	0	.016317
	5	1	.163170
	5	4	.081585
	5	5	.004662
	6	0	.004079
	6	1	.073427
	6	4	.183566
	6	5	.024476
	6	6	.000583
	7	0	.000583
	7	1	.024476
	7	2	.183566
	7	5	.073427
	7	6	.004079
	8	1	.004662
	8	2	.081585
	8	5	.163170
	8	6	.016317
	9	2	.020979
	9	3	.195804
	9	6	.048951
	10	3	.069930
	10	6	.122378
	11	4	.192308

N	n	x	f(x)
14	2	2	.164835
	3	0	.153846
	3	3	.054945
	4	0	.069930
	4	3	.159840
	4	4	.014985
	5	0	.027972
	5	4	.059940
	5	5	.002997
	6	0	.009324
	6	1	.111888
	6	4	.139860
	6	5	.015984
	6	6	.000333
	7	0	.002331
	7	1	.048951
	7	5	.048951
	7	6	.002331
	8	0	.000333
	8	1	.015984
	8	2	.139860
	8	5	.111888
	8	6	.009324
	9	1	.002997
	9	2	.059940
	9	6	.027972
	10	2	.014985
	10	3	.159840
	10	6	.069930
	11	3	.054945
	11	6	.153846
	12	4	.164835
15	2	2	.142857
	3	0	.184615
	3	3	.043956
	4	0	.092308
	4	3	.131868
	4	4	.010989
	5	0	.041958
	5	4	.044955
	5	5	.001998
	6	0	.016783
	6	1	.151049
	6	4	.107892
	6	5	.010789
	6	6	.000200
	7	0	.005594
	7	1	.078322
	7	4	.195804
	7	5	.033566
	7	6	.001399

Table 4. (continued)

N	n	x	f(x)
15	8	0	.001399
	8	1	.033566
	8	2	.195804
	8	5	.078322
	8	6	.005594
	9	0	.000200
	9	1	.010789
	9	2	.107892
	9	5	.151049
	9	6	.016783
	10	1	.001998
	10	2	.044955
	10	6	.041958
	11	2	.010989
	11	3	.131868
	11	6	.092308
	12	3	.043956
	12	6	.184615
	13	4	.142857
16	2	2	.125000
	3	3	.035714
	4	0	.115385
	4	3	.109890
	4	4	.008242
	5	0	.057692
	5	4	.034341
	5	5	.001374
	6	0	.026224
	6	1	.188811
	6	4	.084291
	6	5	.007493
	6	6	.000125
	7	0	.010490
	7	1	.110140
	7	4	.157343
	7	5	.023601
	7	6	.000874
	8	0	.003496
	8	1	.055944
	8	5	.055944
	8	6	.003496
	9	0	.000874
	9	1	.023601
	9	2	.157343
	9	5	.110140
	9	6	.010490
	10	0	.000125
	10	1	.007493
	10	2	.084291
	10	5	.188811
	10	6	.026224

N	n	x	f(x)
16	11	1	.001374
	11	2	.034341
	11	6	.057692
	12	2	.008242
	12	3	.109890
	12	6	.115385
	13	3	.035714
	14	4	.125000
17	2	2	.110294
	3	3	.029412
	4	0	.138655
	4	3	.092437
	4	4	.006303
	5	0	.074661
	5	3	.177763
	5	4	.026665
	5	5	.000970
	6	0	.037330
	6	4	.066661
	6	5	.005333
	6	6	.000081
	7	0	.016968
	7	1	.142534
	7	4	.127262
	7	5	.016968
	7	6	.000566
	8	0	.006787
	8	1	.081448
	8	5	.040724
	8	6	.002262
	9	0	.002262
	9	1	.040724
	9	5	.081448
	9	6	.006787
	10	0	.000566
	10	1	.016968
	10	2	.127262
	10	5	.142534
	10	6	.016968
	11	0	.000081
	11	1	.005333
	11	2	.066661
	11	6	.037330
	12	1	.000970
	12	2	.026665
	12	3	.177763
	12	6	.074661
	13	2	.006303
	13	3	.092437
	13	6	.138655
	14	3	.029412

We enter Table 4 under N = 16, n = 6, x_c = 4 and find P(X=4; 16,6,6) = 0.084291. In order to find P*, we sum up all probabilities P(X=x; 16,6,6) $\leq$ 0.084291:

$$P^* = [P(X=4) = 0.084291] + [P(X=5) = 0.007493] + [P(X=6) = 0.000125] + [P(X=0) = 0.026224] = 0.118133.$$

Since $P^* < \alpha$ and $\hat{p}=4/6 > 6/16=p_o$, line 241 has to be classified as homodyne - a result which corresponds to that obtained in Example 8.

3. *Approximations*

So far we have only considered the use of the HPD in analyzing *single* hexameter lines. If we want to analyze longer passages of verse, however, the values for N might become rather large. Since in the LIEBERMAN/OWEN tables all possible probability distributions are given only for N $\leq$ 50, approximations to the HPD may be very useful. We will therefore discuss at least some of the various existing approximations (for further details see e.g. LIEBERMAN/OWEN 1961; JOHNSON/KOTZ 1969: 148ff and MOLENAAR 1973).

3.1. *Binomial Approximation*

As can be shown (cf. HALD 1952: 42) for N $\rightarrow \infty$ and a fixed value of p = k/N, the HPD converges to the binomial distribution. If we denote the cumulative HPD by H(x;N,n,k) and the cumulative binomial distribution by

$$B(x;n,p) = \sum_{x=0}^{x_c} \binom{n}{x} p^x (1-p)^{n-x},$$

using the identities (7), (8), and (9), we arrive at the following possible approximations (cf. LIEBERMAN/OWEN 1961: 17):

$$H(x;N,n,k) \cong B(x;n,k/N) \quad (15a)$$
$$H(x;N,n,k) \cong B(x;k,n/N) \quad (15b)$$
$$H(x;N,n,k) \cong 1 - B(k-x-1;N-n,k/N) \quad (15c)$$
$$H(x;N,n,k) \cong 1 - B(n-x-1;N-k,n/N) \quad (15d)$$

In some tables of the binomial distribution $P(X \geq x_c)$ is tabulated instead of $P(X \leq x_c)$. If we denote the former probability by E(x;n,k/N), the approximation formulas (15a) to (15d) can also be written as

$$H(x;N,n,k) \cong 1 - E(x+1;n,k/N) \quad (16a)$$
$$H(x;N,n,k) \cong 1 - E(x+1;k,n/N) \quad (16b)$$
$$H(x;N,n,k) \cong E(k-x;N-n,k/N) \quad (16c)$$
$$H(x;N,n,k) \cong E(n-x;N-k,n/N) \quad (16d)$$

From these approximations, that approximation should always be used which is based on the smallest sample size or, in other words, the one which has the smallest value of b in B(a;b,c) or E(a;b,c).

Example 10. Suppose H(3;50,25,5) is to be approximated. The exact value is 0.825662. Using tables of the cumulative binomial distribution (for extensive tables see e.g. NATIONAL BUREAU OF STANDARDS 1950; ROMIG 1953; HARVARD UNIVERSITY COMPUTATION LABORATORY 1955; WEINTRAUB 1963) we obtain the following approximations:

$$H(3;50,25,5) \cong B(3;25,0.1) = 0.7636$$
$$H(3;50,25,5) \cong B(3;5,0.5) = 0.8125$$

$$H(3;50,25,5) \cong 1 - B(1;25,0.1) = 0.7288$$

$$H(3;50,25,5) \cong 1 - B(21;45,0.5) = 0.6170$$

There are large differences among these 4 approximations. The best approximation is $B(3;5,0.5)$, which is the one with the smallest sample size for the binomial distribution. This can be explained by the fact that for $b \leq 0.1N$ there is not much difference between drawing a sample from a finite population, as is the case with the hypergeometric model, or from an infinite population - the latter being the model on which the binomial distribution is based.

It is usually assumed that the above approximations are sufficiently good if the sample size is not larger than, say, 10% of the population, i.e. if $b \leq 0.1N$ in $B(a;b,c)$ or $E(a;b,c)$. It has, however, been shown (cf. MOLENAAR 1973: 142) that at least for the case that $b = 0.1N$ as in the example just discussed, the approximations (15) and (16) overestimate the tails of the HPD and thus may lead to conservative decisions in the case of significance tests.

The above approximations can however be improved if we make the estimation of the binomial parameter p not only dependent on N, n and k but also on x. There exist a number of such improved binomial approximations. We mention only the following one (from WISE 1954), which is fairly simple and sufficiently exact for many applications, provided that n is less than say 0.3N:

$$H(x;N,n,k) \cong B(x;n,p) = \sum_{x=0}^{x_c} \binom{n}{x} p^x (1 - p)^{n-x} \tag{17}$$

with

$$p = \frac{2k - x}{2N - n + 1} \quad \text{and} \quad n \leq k \leq \frac{1}{2}N.$$

It should be mentioned that the condition $n \leq k \leq (1/2)N$ does not result in any loss of generality in the application of (17), since by using the various identities given in Part 1 of this article we can always comply with this convention.

Example 11. Let us assume we want to approximate $H(7;100,42,27) = H(7;100,27,42) = 0.038230$. With the help of (17) we calculate

$$p = \frac{2(42) - 7}{2(100) - 27 + 1} = 0.4425$$

and

$$B(7;27,0.4425) = \sum_{x=0}^{7} \binom{27}{x} 0.4425^x 0.5575^{27-x} = 0.040013$$

which is close to the exact probability, although n is considerably larger than 0.1N.

3.2. Normal Approximation

With the help of the formulas for the mean and the variance (cf. Formulas 5 and 6) the HPD can also be approximated by the normal distribution by means of the transformation

$$z = \frac{X - E(X)}{\sqrt{V(X)}} = \frac{X - \frac{nk}{N}}{\sqrt{\frac{nk(N-n)(N-k)}{N^2(N-1)}}} = \frac{X - np}{\sqrt{npq\frac{N-n}{N-1}}} \tag{18}$$

where $p = k/N$ and $q = 1 - p$. As with the binomial approximation, n and k may be interchanged.

As suggested for instance by LIEBERMAN/OWEN (1951: 21) the normal approximation may be used if

$$V(X) = npq \frac{N - n}{N - 1} > 9$$

which is the case only for relatively large N, and k/N or n/N not too far from 0.5. This can be explained by the fact that the normal distribution is symmetric and that for k/N or n/N approaching 0.5 the HPD tends to be symmetric as well.

Example 12. For H(48;200,100,100) the exact probability is 0.335742. Since

$$V(X) = 100 \frac{1}{2} \frac{1}{2} \frac{100}{199} = 12.56 > 9,$$

we may use the normal approximation which yields

$$z = \frac{48 - 50}{\sqrt{12.56}} = -0.5643.$$

Using tables of the normal distribution we get for $P(Z \leq -0.5643)$ a value of 0.2863, which is not a good approximation to the exact probability.

The normal approximation (18) can,however,be improved by using the so-called correction for continuity, which takes into account that the normal distribution is continuous, the HPD however discrete. The normal approximation corrected for continuity is

$$Z = \frac{X + 0.5 - E(X)}{\sqrt{V(X)}} = \frac{X + 0.5 - \frac{nk}{N}}{\sqrt{\frac{nk(N-n)(N-k)}{N^2(N-1)}}} = \frac{X + 0.5 - np}{\sqrt{npq \frac{N-n}{N-1}}}. \quad (19)$$

For H(48;200,100,100) we now get $P(Z \leq -0.4243) = 0.3361$ which is close to the exact probability.

As the example shows, the correction for continuity may lead to a considerable increase in accuracy. Therefore, unless $V(X) >> 9$ we should always use (19) instead of (18). It should be noted that (19) may be sufficiently accurate even if $V(X) < 9$.

As with the binomial distribution there exist several improved normal approximations which are based not only on N, n and k but also on x. We will only present the following square root approximation which is very simple and especially accurate

near the customary significance levels (cf. MOLENAAR 1973: 132f):

$$Z = \frac{2}{\sqrt{N-1}} [\sqrt{(x+1)(N-n-k+x+1)} - \sqrt{(n-x)(k-x)}] \quad (20)$$

for $n \leq k \leq \frac{1}{2}N$.

Example 13. We now want to approximate H(5;50,20,20) = = 0.069248. Note that $V(X) = 2.9388$ and $n = 0.4N$ and that therefore the condition for the application of the normal approximations (18) and the condition for the application of the binomial approximations (15) and (16) are not fulfilled. The square root normal approximation yields

$$z = \frac{2}{\sqrt{50-1}} [\sqrt{(5+1)(50-20-20+5+1)} - \sqrt{(20-5)(20-5)}] = -1.4863.$$

The corresponding probability is 0.068600,which is a good approximation to the exact value.

We now use the approximations (17), (18) and (19) and obtain the following probabilities:

$$B(5;20,0.4321) = 0.075553$$
$$P(Z \leq -1.75) = 0.040059$$
$$P(Z \leq -1.46) = 0.072374.$$

As could be expected, (18) is very inaccurate. (19), however, is relatively close to the exact value, although $V(X) << 9$,and even (17) leads to a good approximation,although $n > 0.3N$.

3.3. Chi-Square Approximation

As can be shown, the statistic

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}, \tag{21}$$

where O_i designates the observed and E_i the expected values, is asymptotically distributed as χ^2 with k-1 degrees of freedom (df). In the case of a 2x2 table such as Table 2 the corresponding χ^2 has one degree of freedom, and (21) can be written as

$$\chi^2 = \frac{N(ad - bc)^2}{knrs}, \tag{22}$$

or with the correction for continuity as

$$\chi^2 = \frac{N(|ad - bc| - 0.5N)^2}{knrs}. \tag{23}$$

(22) and (23) can be considered as the classical approximations to the HPD in the analysis of contingency tables. They may be applied if $N \geq 20$ and all expected values ≥ 5. This is, however, only a rule of thumb which is not generally accepted. Furthermore, it is advisable to always use the formula corrected for continuity unless N is very large.

Since b = n-a, c = k-a and d = N-n-k+a, (22) and (23) can also be written as

$$\chi^2 = \frac{(a - \frac{nk}{N})^2 N^3}{knrs} \tag{24}$$

and

$$\chi^2 = \frac{(|a - \frac{nk}{N}| - 0.5)^2 N^3}{knrs}. \tag{25}$$

Since, for df = 1, $z^2 = \chi^2$, a further χ^2 approximation can be derived by squaring the normal approximation (18). Setting r = N-n, s = N-k and a = X we obtain from (18)

$$\chi^2 = \frac{(a - \frac{nk}{N})^2 N^2 (N-1)}{knrs}. \tag{26}$$

If we now take into account that the χ^2 test is a two-tailed test and that in this case the normal variate corrected for continuity is

$$z = \frac{|X - E(X)| - 0.5}{\sqrt{V(X)}},$$

we get from (26) the following χ^2 statistic corrected for continuity which is equivalent to the normal approximation (19) for $|a-nk/N| \geq 0.5$:

$$\chi^2 = \frac{(|a - \frac{nk}{N}| - 0.5)^2 N^2 (N-1)}{knrs}. \tag{27}$$

It is easily seen that (26) and (27) are asymptotically equivalent to (24) and (25) and thus also to (22) and (23) and that therefore the difference between these test statistics becomes negligible for sufficiently large N. It is sometimes argued that (26) and (27) are preferable to (22) and (23) for smaller values of N. But as MOLENAAR (1973: 130) has shown, the classical χ^2 approximation (23) is in most cases better than (27). Therefore (23) and its equivalent (25) should be prefered to (27) as an approximation to the HPD.

Example 14. Suppose we had observed the following 2x2 table:

3	11	14
9	7	16
12	18	30

The exact hypergeometric probability $P(X \leq 3; 30, 14, 12)$ is 0.057298. The X^2 approximation (23) yields

$$X^2 = \frac{[|3(7) - 11(9)| - 15]^2 30}{12(14)16(18)} = 2.4609.$$

In order to compare the X^2 value with the exact hypergeometric probability we compute

$$z_{\alpha/2} = \sqrt{X^2_\alpha} = \sqrt{2.4609} = 1.5687.$$

We find that $P(Z \geq 1.5687) = 0.058355$, which is close to the exact probability although N is not very large.

4. Analysis of Longer Passages of Verse

We now want to show how to use the approximations discussed in the previous chapter in the analysis of longer passages of verse. We assume in this context that the conditions for the application of the different approximations are fulfilled. Otherwise one should proceed as in the analysis of single hexameter lines.

As has been shown by MOLENAAR (1973), the square root normal approximation (20) and in particular the binomial approximation (17) are in many cases better than the classical X^2 approximation (23). Since furthermore both (17) and (20) are also very easy to compute we suggest using for a one-tailed test (17) or (20) and then classifying according to (14).

Since the X^2 test in contradistinction to (17) and (20) leads directly to two-tailed probabilities and is furthermore simpler to apply if $\hat{p} > p_o$, we suggest using the X^2 statistic (23) for a two-tailed test. We then classify as follows:

$$\begin{array}{lll} \text{heterodyne:} & P(\chi^2 \geq X^2) \leq \alpha & \text{and } \hat{p} < p_o \\ \text{homodyne:} & P(\chi^2 \geq X^2) \leq \alpha & \text{and } \hat{p} > p_o \\ \text{neutral:} & P(\chi^2 \geq X^2) > \alpha & \end{array} \qquad (28)$$

Furthermore, since both the one-tailed test and the two-tailed test become more and more powerful when N increases, for the analysis of longer passages of verse we suggest choosing values of α which are smaller than those chosen for the analysis of single hexameter lines.

Example 15. Table 5 shows the relation between ictus and accent in *Iliad* 1, 237-239.

Table 5. Relation between Ictus and Accent in *Iliad* 1, 237-239

Syllables	heavy	light	
accented	7	10	17
unaccented	11	21	32
	18	31	49

We carry out a two-tailed test with $\alpha = 0.10$. Using (23) we obtain $X^2 = 0.02522$. Since $P(\chi^2 \geq 0.02522) = 0.873820 > \alpha = 0.10$ the three lines have to be classified as neutral.

Example 16. Table 6 shows the relation between ictus and accent in *Iliad* 1, 304-311 and *Odyssey* 4, 842-847. Both passages are classified by SCHMIEL (in this volume p. 16ff) as heterodyne and furthermore as narrative and calm.

Table 6. Relation between Ictus and Accent in *Iliad* 1, 304-311 and *Odyssey* 4, 842-847

Syllables	*Iliad* heavy	*Iliad* light		*Odyssey* heavy	*Odyssey* light	
accented	11	23	34	6	22	28
unaccented	37	57	94	30	37	67
	48	80	128	36	59	95

Since according to SCHMIEL narrative and calm passages tend to be heterodyne, we carry out a one-tailed test with $\alpha = 0.05$. We use (17) because for both texts $n < 0.3N$. In the case of the *Iliad* we have to approximate $H(11;128,34,48)$. The binomial approximation (17) yields

$$B(11;\ 34,\ 0.3812) = 0.307163.$$

Analogously we obtain for the passage from the *Odyssey*

$$H(6;\ 95,\ 28,\ 36) \cong B(6;\ 28,\ 0.4049) = 0.027973.$$

Only the value for the passage taken from the *Odyssey* is less than α. Therefore SCHMIEL's hypothesis of heterodyne texture is confirmed only for the lines from the *Odyssey*.

If we had applied the square root normal approximation (20) to the data of Table 6, we would have obtained for the *Iliad* $P(Z \leq -0.49516) = 0.310243$ and for the *Odyssey* $P(Z \leq$ $\leq -1.93514) = 0.026486$, which is essentially the same result as that obtained with the help of (17).

Considering the above result, I think it would be interesting to reanalyze all the data presented by SCHMIEL in order to see whether his hypothesis that there is a correlation between *form* (homodyne, heterodyne, and neutral) and *content* (calm, narrative etc.) can be maintained when we use the classification procedure proposed in the present paper.

5. *Measuring the Strength of the Relation between Ictus and Accent*

5.1. *The φ coefficient*

With the help of the HPD and the various approximations discussed above we can test whether the observed number of coincidences deviates significantly from the number expected under H_0 or whether a deviation observed should be attributed to chance. In carrying out such a test we take into account

the magnitude of the difference between the observed and the expected number of coincidences as well as the magnitude of the sampling error, which is a function of N, n and k. Let us illustrate this by means of Table 7.

Table 7. Two 2x2 tables both with $\hat{p} = 0.33$ and $p_0 = 0.4$

(a)

10	20	30
8	7	15
18	27	45

(b)

20	40	60
16	14	30
36	54	90

In Table 7a as well as in Table 7b $\hat{p} = 0.33$ and $p_0 = 0.4$. Since $\hat{p} < p_0$, the relation is negative although $a > c$. It is furthermore relatively weak because $\hat{p}$ and p_0 do not differ very much. As for both tables $\hat{p} = 0.33$, $p_0 = 0.4$, the *strength* of the relation between ictus and accent is the same in both cases. Since, however, all values in Table 7b are twice as large as in Table 7a, the sampling error is smaller in the case of Table 7b, and thus a test of significance will yield a smaller probability for this table. This is immediately seen if we compute $H(x;\ N,n,k)$ for both tables. We get $H(10;\ 45,30,18) =$ $= 0.166438$ and $H(20;\ 90,60,36) = 0.055545$ and thus a smaller value for Table 7b.

If we carry out a X^2, test the dependence of the X^2 value on the magnitude of N, n and k becomes even more clear. With the help of (22) we obtain for Table 7a $X^2 = 1.67$ and for Table 7b $X^2 = 3.34$, the latter value being exactly twice as large as the former one. If we had multiplied all the values of Table 7a not by 2 but by, say, 10 we would have obtained a X^2 value ten times as large as that of Table 7a and consequently a very low probability, although the difference between $\hat{p}$ and p_0 would not have been affected by this multiplication. Hence, interpreting a low probability in significance tests as an indication of a strong relationship between two (or more) variables may be very misleading.

Thus, in order to be able to compare the strength of the relationship in different texts, we need a measure which is not dependent on the sampling error. There exist a vast number of coefficients meeting this criterion. One of the coefficients most often used in the case of 2x2 tables is the φ coefficient. There exist several possibilities to derive the φ coefficient. It can, e.g., be obtained from the well known product-moment correlation coefficient

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{N\Sigma xy - \Sigma x\Sigma y}{\sqrt{[N\Sigma x^2 - (\Sigma x)^2][N\Sigma y^2 - (\Sigma y)^2]}} \qquad (29)$$

as follows:
Let us consider ictus and accent as two dichotomous variables X and Y which both can assume the values 1 and 0. In this case the relation between ictus and accent can be represented as shown in Table 8.

Table 8. Relation between Ictus (X) and Accent (Y)

		X heavy 1	X light 0	
Y	accented 1	a	b	a+b
Y	unaccented 0	c	d	c+d
		a+c	b+d	N

If we now take into account that $N = a+b+c+d$, $\Sigma xy = a$, $\Sigma x = \Sigma x^2 = a+c$, and $\Sigma y = \Sigma y^2 = a+b$, (29) can be written after some elementary rearrangements as

$$\varphi = \frac{ad - bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}} \, . \qquad (30)$$

Since φ is derived from r_{xy}, it has the same range as r_{xy}, namely $-1 \leq \varphi \leq 1$. A negative value of φ indicates that $\hat{p} < p_o$ and a positive one that $\hat{p} > p_o$. It should however be noted that this interpretation of the sign of φ is only true if the variables are coded as in Table 8. If we code e.g. 'light' as 1, a negative value of φ would indicate that $\hat{p} > p_o$ and a positive one that $\hat{p} < p_o$.

If we compare the formula for φ with the X^2 statistic (22), we immediately realize that φ can also be expressed as

$$\varphi = \sqrt{\frac{X^2}{N}} \, . \qquad (31)$$

Since X^2 is approximately distributed as χ^2 with df = 1, we can test the φ coefficient for significance, provided that the conditions for the application of (22) are fulfilled, simply by computing

$$X^2 = N\varphi^2 \qquad (32)$$

or, since for df = 1 $\sqrt{\chi^2_\alpha} = z_{\alpha/2}$, by calculating

$$z_{\alpha/2} = \varphi\sqrt{N}. \qquad (33)$$

We now compute the φ coefficient for Tables 7a and 7b. We obtain in both cases $\varphi = -0.1925$. As could be expected, φ is negative, since $\hat{p} < p_o$. Furthermore, the value of φ is the same for both tables, because φ measures the strength of a relation independently from the sampling error.

The point is now how to interpret a φ coefficient of -0.1925 in terms of the strength of the relationship it measures.[4] If we consider the marginals of a 2x2 table as random variables, φ can theoretically assume all the values within the interval [-1;1]. In this case φ has to be interpreted with respect to -1 or +1 being the maximum values theoretically possible. If we consider the marginals of a given table as fixed, however, the maximum theoretically possible value of φ may be only -0.7 or even only 0.5. In this case we have to interpret φ not with respect to $\pm$ 1 but with respect to its theoretically possible maximum given the marginals of the table under investigation. We thus have to determine the maximum φ for tables with fixed marginals (henceforward φ_{max}).

5.2. *A Modified φ Coefficient for 2x2 Tables with Fixed Marginals*

It is immediately obvious that there is a maximum of positive association between X and Y if all syllables which have the value 1 in X also have the value 1 in Y and vice versa. This is only possible, however, if a+b = a+c *and* b = d = 0. It then follows immediately from (30) that φ_{max} = 1. Analogously, there is a maximum of negative association if a+b = b+d *and* a = d = 0. Then φ_{max} = -1. If however neither a+b = a+c nor a+b = b+d, only one of the four cells of a 2x2 table can assume the value 0. In that case we determine φ_{max} as follows: since in the case of positive association, i.e. if $\hat{p} > p_o$ or equivalently ad-bc > 0, φ assumes its maximum if both a and d assume their maximum, we first determine max(a) and max(d). According to (3) max(a) = min(a+b;a+c) and max(d) = min(c+d;b+d). Furthermore, since b or c must be equal to zero if both a and d assume their maximum, it follows directly from (30) that

$$\varphi_{max} = \frac{\min(a+b;a+c)\cdot\min(c+d;b+d)}{\sqrt{(a+b)(a+c)(b+d)(c+d)}} . \qquad (34a)$$

In the case of negative association, i.e. if ad-bc < 0, we have to determine max(b) and max(c). We get max(b) = min(a+b;b+d) and max(c) = min(a+c;c+d). Since this time a *or* d must be 0, it follows from (30) that

$$\varphi_{max} = -\frac{\min(a+b;b+d)\cdot\min(a+c;c+d)}{\sqrt{(a+b)(a+c)(b+d)(c+d)}} . \qquad (34b)$$

We now compute (34a) and (34b) for Table 7. We obtain for Table 7a as maximum positive association

$$\varphi_{max} = \frac{18(15)}{\sqrt{18(27)30(15)}} = 0.5774$$

and as maximum negative association

$$\varphi_{max} = -\frac{27(15)}{\sqrt{18(27)30(15)}} = -0.8660.$$

We now calculate φ_{max} for Table 7b and get the same values, as could be expected.

As the example shows, φ_{max} can be very different from -1 or +1. Furthermore, the range of φ is in many cases not symmetric about zero.

With the help of (34) we can now define a modified φ coefficient for 2x2 tables with fixed marginals as

$$\varphi' = \frac{\varphi}{|\varphi_{max}|} , \qquad (35)$$

where φ_{max} is given by (34a) if φ is positive and by (34b) if φ is negative. In contradistinction to φ, φ' is symmetric about zero and can attain the values -1 and +1 for any set of fixed marginals.

For the data of Table 7 (35) yields

$$\varphi' = -\frac{0.1925}{0.8660} = 0.2223,$$

a value which is only slightly higher than the original φ. However, if we had observed a value of +0.1925 for the data of Table 7, we would have obtained φ = 0.1925/0.5774 = 0.3334, which is considerably higher than the original φ.

Since the denominators of (30) and (34) are identical, the division of φ by φ_{max} yields in the case of positive or zero association

$$\varphi' = \frac{ad-bc}{\min(a+b;a+c)\cdot\min(c+d;b+d)}\,, \quad 0 \leq \varphi \leq 1 \qquad (36a)$$

for ad - bc > 0

and in the case of negative or zero association

$$\varphi' = \frac{ad - bc}{\min(a+b;b+d)\cdot\min(a+c;c+d)}\,, \quad -1 \leq \varphi \leq 0 \qquad (36b)$$

for ad - bc < 0.

These two formulas, which are very simple to calculate, can be used if one wants to compute φ' without previous computation of φ_{max}.[5] Since $\varphi = \varphi'\cdot|\varphi_{max}|$ it follows directly from (33) that φ' can be tested for significance by using

$$z_{\alpha/2} = \varphi'\cdot|\varphi_{max}|\sqrt{N}. \qquad (37)$$

<u>Example 17</u>. The use of (36) is illustrated by the data of Example 16. With the help of (36b) we obtain for the *Iliad*

$$\varphi' = \frac{11(57) - 23(37)}{34(48)} = -0.1373$$

and for the *Odyssey*

$$\varphi' = \frac{6(37) - 22(30)}{28(36)} = -0.4345.$$

Compared to $\varphi'_{max} = -1$, only the value for the *Odyssey* indicates a relatively high degree of heterodyne structure.

If one wants to test whether there is a significant difference between the strength of the relation between ictus and accent in two longer passages of verse, one can proceed as fol-

lows: we first calculate φ' for both texts. Since $\varphi = Z/\sqrt{N}$, $E(Z) = 0$ and $V(Z) = 1$, then

$$E(\varphi') = E\left(\frac{\varphi}{|\varphi_{max}|}\right) = \frac{1}{|\varphi_{max}|}E\left(\frac{Z}{\sqrt{N}}\right) = 0 \qquad (38)$$

and

$$V(\varphi') = V\left(\frac{\varphi}{|\varphi_{max}|}\right) = \frac{1}{\varphi^2_{max}}\,V\left(\frac{Z}{\sqrt{N}}\right) = \frac{1}{N\varphi^2_{max}}\,. \qquad (39)$$

We now transform the difference $\varphi'_1 - \varphi'_2$ into a normal variable

$$Z = \frac{\varphi'_1 - \varphi'_2 - E(\varphi'_1 - \varphi'_2)}{\sqrt{V(\varphi'_1) + V(\varphi'_2)}}$$

and arrive at the test statistic

$$Z = \frac{\varphi'_1 - \varphi'_2}{\sqrt{\dfrac{1}{N_1\varphi^2_{1max}} + \dfrac{1}{N_2\varphi^2_{2max}}}}\,, \qquad (40)$$

which is approximately normally distributed with zero mean and unit variance.

<u>Example 18</u>. Let us carry out a two-tailed test with $\alpha = 0.10$ for the data of Example 16. (40) yields

$$z = \frac{-0.1373 + 0.4345}{\sqrt{\dfrac{1}{128(-0.4634^2)} + \dfrac{1}{95(-0.5050^2)}}} = 1.0667.$$

Since $P(Z \geq 1.0667) = 0.1430 > \alpha/2 = 0.05$ the difference between the values of φ' in both texts can be regarded as non-significant. Thus the hypothesis of a statistical interaction between the variables 'number of coincidences' and 'kind of text' (the *Odyssey* vs. the *Iliad*) is not confirmed.

With regard to this interpretation of the test result it should however be noted that the value of the test statistic (40) is largely a function of both the N_i involved in the computation of the variance. Thus, for instance, if one N_i is very small, this N_i may have a strong influence on the numerical value of the variance. Furthermore, since the value for the variance may become relatively large in this case, the test result may be non-significant. Therefore, in order to attain maximum efficiency, both N_i should have approximately the same size.

6. Conclusion

As has already been mentioned, the statistical model for the coincidence of ictus and accent presented in this paper is not restricted to the analysis of hexameter verse but can also be applied to other versification systems. Future research should therefore not only analyze various kinds of hexameter texts, but should also extend the application of the model to other types of verse.

Furthermore, there are numerous, rather sophisticated methods which have been developed for the analysis of contingency tables (see e.g. the references given by BLALOCK 1979: 332f). With the help of these methods one can not only analyze first-order interactions such as between 'number of coincidences' and 'kind of text' but also higher-order interactions of various kinds. The strength of the relationship between the variables in such multi-dimensional contingency tables may be measured by a number of coefficients, some of which also permit the measurement of partial associations.

Summary

In the present article by means of some combinatorial deliberations we first derive the hypergeometric probability distribution as a model for the analysis of the coincidence of ictus and accent. On the basis of this model a table is developed which permits an objective classification of hexameter verse according to the number of coincidences of ictus and accent. Subsequently the use of some approximations to the hypergeometric model in the analysis of longer passages of verse is discussed. Finally a coefficient is presented which allows one to measure the strength of association between ictus and accent. In order to facilitate the understanding of the statistical procedures discussed in this article all procedures are illustrated by at least one application.

Notes

1 Three different types of sampling situations leading to the following three types of 2x2 tables can be differentiated: (1) tables with both margins fixed (2) with one margin fixed and (3) with no margin fixed (cf. the discussion in KENDALL/STUART 1973: 569ff). Although in the present case one might think that only one margin should be considered as fixed (i.e. the number of heavy and light syllables), in the light of the sampling situation involved it seems in my opinion appropriate to consider both margins as fixed. As has been shown by TOCHER (1950), the HPD can be used in the analysis of 2x2 tables of type (2) and (3) as well.

2 TOCHER (1950) has however shown that a slight modification of the one-tailed Fisher-Yates test which consists of introducing an additional random process makes this test a uniformly most powerful unbiassed test for 2x2 tables with both margins, one margin or no margin fixed.

3 For a discussion of different definitions of rejection regions for the two-tailed Fisher-Yates test see ARMSEN (1955).

4 Unfortunately φ does not have such a clear-cut interpretation as e.g. the so-called *proportional reduction in error measures* (for a critique of φ and other χ^2-based contingency coefficients see COSTNER 1965 and BLALOCK 1979: 299ff).

5 For a somewhat different approach to defining φ' see JÄGER (1976), GEBERT (1977) and BÖSSER (1979).

References

ARMSEN, P., Tables for significance tests of 2x2 contingency tables. Biometrika 42, 1955, 494-511

BENNETT, B.M., HORST, C., Suppelement to Tables for Testing Significance in a 2x2 Contingency Table. Cambridge, Cambridge University Press 1966

BLALOCK, H.M.Jr., Social Statistics. Rev. 2nd edition. Tokyo, McGraw-Hill Kogakusha 1979

BÖSSER, T., PHI-Interkorrelationen seltener Symptome. Psychologische Beiträge 21, 1979, 343-348

COSTNER, H.L., Criteria for Measures of Association. American Sociological Review 30, 1965, 341-353

DEFENSE SYSTEMS DEPARTMENT, GENERAL ELECTRIC COMPANY, Tables of the Individual and Cumulative Terms of Poisson Distribution. Princeton, N.J. 1962

FELDMAN, S.E., KLINGER, E., Short cut calculation of the Fisher-Yates "exact test". Psychometrika 28, 1963, 289-291

FINNEY, D.J., The Fisher-Yates Test of Significance in 2x2 Contingency Tables. Biometrika 35, 1948, 145-156

FINNEY, D.J., LATSCHA, R., BENNETT, B.M., HSU, P., Tables for Testing Significance in a 2x2 Contingency Table. Cambridge, Cambridge University Press 1963

FISHER, R.A., The logic of inductive inference. Journal of the Royal Statistical Society 98, 1935, 39-54

FREEMAN, G.H., HALTON, J.H., Note on an exact treatment of contingency, goodness of fit and other problems of significance. Biometrika 38, 1951, 141-149

GEBERT, A., JÄGERs Phi(G) als Item-Interkorrelationsmaß für Faktorenanalysen. Psychologische Beiträge 19, 1977, 336-339

HALD, A., Statistical Theory with Engineering Applications. New York, Wiley 1952

HARVARD UNIVERSITY COMPUTATION LABORATORY, Tables of the Cumulative Binomial Probability Distribution. Cambridge, Mass., Harvard University Press 1955

JÄGER, R., φ_G - ein statistischer Test zur Bestimmung der Differenzierungsfähigkeit psychologischer Skalen. Psychologische Beiträge 18, 1976, 214-223

JOHNSON, F.M., The Fisher-Yates exact test and unequal sample sizes. Psychometrika 37, 1972, 103-106

JOHNSON, N.L., KOTZ, S., Distributions in Statistics. Discrete Distributions. Boston, Houghton Mifflin Company 1969

KENDALL, M.G., STUART, A., The Advanced Theory of Statistics. Vol. 2. London, Griffin 1973

LIEBERMAN, G.J., OWEN, D.B., Tables of the Hypergeometric Probability Distribution. Stanford, Cal., Stanford University Press 1961

LIENERT, G.A., Verteilungsfreie Methoden in der Biostatistik. Tafelband. Meisenheim, Anton Hain 1975

MOLENAAR, W., Approximations to the Poisson, Binomial and Hypergeometric Distribution Functions. Amsterdam, Mathematisch Centrum 1973

MOLINA, E.C., Poisson's Exponential Binomial Limit. New York, Van Norstrand 1945

MOOD, A.M., GRAYBILL, F.A., BOES, D.C., Introduction to the Theory of Statistics. Tokyo, McGraw-Hill Kogakusha 1974

NATIONAL BUREAU OF STANDARDS, Tables of the Binomial Probability Distribution, Washington 1950 (Appl. Math. Series No. 6)

ROMIG, H.G., 50-100 Binomial Tables. New York, Wiley 1953

SIEGEL, S., Nonparametric Statistics for the Behavioral Sciences. New York, McGraw-Hill 1956

TOCHER, K.D., Extension of the Neyman-Pearson theory of tests to discontinous variates. Biometrika 37, 1950, 130-144

WALL, K.D., Kombinatorische Analyse von Kontingenztafeln. Berlin, Unpublished doctoral dissertation. Technische Hochschule Berlin 1972

WEINTRAUB, S., Tables of the Cumulative Binomial Distribution for Small Values of p. London, The Free Press of Glencoe, Collier-Macmillan 1963

WISE, M.E., A quickly convergent expansion for cumulative hypergeometric probabilities, direct and inverse. Biometrika 41, 1954, 317-329

YATES, F., Contingency Tables Involving Small Numbers and the χ^2 Test. Supplement to the Journal of the Royal Statistical Society 1, 1934, 2, 217-235

ETUDE STATISTIQUE SUR L'HEXAMETRE D'AUSONE

Pol Tordeur, Renaix.

Abstract

A Statistical Study of the Hexameters of Ausonius

In this publication we are attempting to determine the style of several poems according to the frequency of the different metrical word types. Having applied the χ^2 test to the total of the tables we look for cells with significant values. We then regroup the observed frequencies of equivalent metrical word types in order to find out whether the results obtained by applying the χ^2 test to the total of the tables turn out to be different.

1. *Introduction*

Le but de cette étude est double:

1°) comparer entre elles différentes oeuvres de l'écrivain latin Ausone par une étude statistique ayant pour objet la fréquence de différents schémas métriques de mots.

2°) examiner si les résultats obtenus sont modifiés lorsque l'on regroupe des schémas de mots métriquement équivalents.

1.1. Ausone a beaucoup écrit, tant en prose qu'en vers. Seule la partie poétique de son oeuvre retiendra ici notre attention. Encore fallait-il définir le mètre à analyser, puisque les vers d'Ausone appartiennent à des rythmes différents. Notre choix s'est porté sur une partie de ses poèmes rédigés en hexamètres dactyliques:

a) un groupe que, par convention et par commodité, nous appellerons "négligé": *Ephemerides, Domestica, Parentalia, Professores* et *Eclogae*[1]. Il s'agit de divers poèmes, généralement courts. Total: 1001 vers.

b) un 2e groupe, nommé, pour les mêmes raisons "soigné": *Cupido + Mosella*[2]. Total: 582 vers.

c) les *Epistulae*: 657 hexamètres.

d) les *Epigrammata*: 282 hexamètres. Pour ces deux dernières oeuvres il est évident que seuls les vers complets et ne comportant que des mots latins ont été étudiés.

Et enfin deux ouvrages d'égale longueur, dont on essaiera de déceler s'ils peuvent être assimilés à un des quatre groupes déjà définis:

e) l'*Ordo Vrbium Nobilium*, 168 vers.

f) le *Technopaegnion*, 163 vers. Ce dernier appelle une remarque: il revêt une forme tellement bizarre[3] - Ausone, le cas est unique, en a terminé tous les vers par un monosyllabe - que les calculs ont d'abord été effectués sans en tenir compte (tableaux 1 et 2: Total 2690 vers), puis en l'y incorporant (tableaux 1, dernières lignes, et 3; Total 2853 vers). Les différences observées sont, on le verra, surprenantes.

Au total, donc, 2.853 vers seront analysés, soit environ 85% des hexamètres d'Ausone. Ne sont pas retenus les poèmes trop brefs (*Praefatiunculae, Griphus, De XII Caesaribus*), fragmentaires (*Epitaphia, Bissula, De Fastis*) ou d'attribution incertaine (*Appendix Ausoniana e.g.*)[4].

1.2. L'analyse à laquelle nous allons procéder s'applique à la fréquence des différents types de schémas métriques; pour ce faire, nous avons eu recours au test de Pearson ou test χ^2.

On sait que ce test permet, par l'application de la formule

$$\chi^2 = \sum_i \frac{\left(o_i - e_i\right)^2}{e_i}$$

où *e* (*expected*) désigne le nombre d'occurences attendu, prévu et *o* (*observed*) le nombre effectivement constaté, de déterminer la probabilité d'attribuer au seul hasard le phénomène observé. Traditionnellement, on définit le seuil de probabilité à $P = 0{,}05$; on considère donc le phénomène comme "voulu" par l'auteur si les chances de l'imputer au hasard n'excèdent pas 5 %.

S'il est facile d'apprécier l'effectif théorique dans un

tableau restreint, l'opération est toujours longue et fastidieuse lorsque les cases sont aussi nombreuses que dans notre tableau n°1. Aussi confie-t-on habituellement le travail à une calculatrice électronique.

En outre, l'estimation du χ^2 critique (correspondant à $P = 0{,}05$) n'est pas fournie par les tables habituelles pour le nombre de cases aussi élevé que celui que nous obtenons. On recourt alors à la loi normale en utilisant la formule[5]

$$\chi^2 = \frac{1}{2}\left(\sqrt{2n-1} + z\right)^2.$$

1.3. Afin d'écarter les effectifs trop réduits, quelques types peu représentés ont été négligés: les monosyllabes élidés et divers mots longs, dont l'alignement aurait nécessité une vingtaine de colonnes supplémentaires et dont la somme se révèle inférieure à 5 % du total des mots. Dans le même but, les effectifs de quatre types de mots très semblables: ∪ - ∪∪, - - ∪∪, ∪∪ - - et ∪∪ - ∪∪, dont la fin coïncide toujours avec le 4e ou le 5e pied, apparaissent groupés dans la colonne 14.[6]

En plus, dans chaque tableau, on a procédé à l'addition des effectifs des types 9 (∪ - ∪) et 10 (∪ - -): c'est que ceux-ci, localisés en fin de vers essentiellement, sont quasiment interchangeables. On pourra dès lors examiner si le manque d'un type est compensé par le surplus de l'autre type - et la différence d'ensemble sera réduite - ou si, au contraire, l'addition de leurs effectifs continue à refléter une attitude semblable du poète à leur égard.

2. *Confection des Tableaux*

Le tableau 1 comporte les effectifs bruts des différents types prosodiques, *Technopaegnion* exclu, puis, dans les deux dernières lignes, *Technopaegnion* inclus.

Le tableau 2 propose pour l'ensemble des oeuvres considérées à l'exception du *Technopaegnion*, les valeurs du test χ^2 arrondies au centième.

Le tableau 3 remplit le même office que le précédent, mais le *Technopaegnion* y est cette fois incorporé.

Le tableau 4 indique en pourcentage l'importance relative par oeuvre de chaque schéma métrique. On en déduit immédiatement l'ordre de fréquence.

3. *Observation des tableaux n°1 et 4.*

3.1. *Premier total, sans le Technopaegnion.*

Une grande uniformité se dégage de toutes les oeuvres: le type le plus employé est le monosyllabe long (col. 2: -, tableau 1) avec 3065 unités, puis vient le spondée représenté par 2479 occurrences (col. 6: - -), ce dernier étant préféré au premier dans les seuls extraits soignés (2e ligne): c'est vraisemblablement une caractéristique du "style soigné" (cf. TORDEUR 1972 § 2 et *infra*).

Partout le trochée occupe le 3e rang (col. 4, 1712: - ᴗ) précédant le dactyle (col. 8, 1405: - ᴗᴗ) puis l'anapeste (col. 7, 1203: ᴗᴗ -) sauf dans les *Epistulae* où 8 succède à 7, séparé de lui par 4 unités seulement - ce qui, sur 3870 mots, est négligeable.

Les neuf colonnes non encore citées se groupent comme suit: dans l'ensemble, 3 (ᴗᴗ), 5 (ᴗ -), 12 (- - -) et 14, en ordre variable, précèdent 9 (ᴗ - ᴗ), 10 (ᴗ - -), 11 (- - ᴗ), 13 (- ᴗᴗ -) et 1 (ᴗ). Les monosyllabes brefs occupent donc une situation opposée à celle de leurs homologues longs: le nombre de syllabes n'importe donc pas tant que la durée du mot.

3.2. *Total général.*

Après l'inclusion du *Technopaegnion,* l'ordre de préférence déjà constaté reste inchangé. Au total, l'ordre de 3 et 12 est permuté à l'exclusion de tout autre changement. Mais à l'intérieur de cette nouvelle oeuvre, *Technopaegnion,* l'ordre est bouleversé: si le monosyllabe long garde la tête, en revanche, le trochée précède - et de loin - le spondée, et l'anapeste vient en 4e place, suivi de peu par le pyrrhique (col. 3) et l'iambe (col.5).

Cette modification est imputable au caractère insolite de l'oeuvre, que nous avons déjà évoqué: Ausone y recourt à des fins de vers de type - ᴗ + ᴗ - + - ou éventuellement ᴗᴗ - + - par exemple tandis qu'il exclut le spondée du dernier pied. Dans les autres oeuvres, par contre, la finale constitue précisément une des places de prédilection du spondée.

4. *Observation du tableau n°2.*

4.1. Il importe tout d'abord d'examiner si la valeur du χ^2 obtenue, 227,49, est significative, le seuil choisi étant P = 0,05. Vu le très grand nombre d'occurrences des divers types - près de 16.000 - et comme les cases du tableau où une différence est susceptible de se manifester sont elles aussi nombreuses - septante au total -, on est en droit de s'attendre à ce que le résultat soit effectivement significatif. Et c'est bien le cas puisque la valeur coïncidant à une probabilité de 0,05 - c'est-à-dire de moins de 5 "chances" sur 100 pour que le hasard soit responsable de la répartition observée - est de 68,39.[7]

4.2. Quelques cases apportent une très forte contribution au total. Ce sont:

soigné, type	2,	monos.longs	25,81
Epigr.	2,	monos.longs	16,57
Epigr.	11,	palimbacchées	13,79
soigné,	9,	amphibraques	10,30
Epist.	5,	iambes	8,39
soigné,	10,	bacchées	8,20
négligé	2,	monos.longs	8,06

etc..., et, sur la ligne des totaux, l'on constate tout aussi aisément que le monosyllabe long, par exemple, présente le chiffre le plus élevé: 56,60, soit près du cinquième du total.

4.3. Si des contributions élevées reflètent de façon évidente des différences importantes, il est en revanche plus délicat d'apprécier des écarts plus ténus. C'est ici que le pourcentage pourrait intervenir, avec ses distensions marquées comme avec ses faibles divergences. Ainsi, la répartition en % laisse elle aussi apparaître de réelles distorsions (cf. tableau 4): par exemple, 15,32 % du type 2 dans les oeuvres soignées et 23,35 % dans les *Epigrammes*, alors que la moyenne est de 19,16 %. Aussi va-t-on bientôt calculer par case selon la méthode définie par RAO (1973).

4.4. A présent, penchons-nous sur la partie du tableau 2, où, comme nous l'avons annoncé précédemment (1.2), les effectifs des types 9 et 10 ont été regroupés.

Le χ^2 est cette fois encore significatif, même s'il ne s'élève plus qu'à 201,01 au lieu de 227,49. Cette diminution reflète normalement la diminution du nombre des cases où se manifestent des variations et montre parfaitement qu'un surplus d'amphibraques (type 9) est compensé par un manque de bacchées (type 10) et vice versa.

4.5. Les cases qui apportent une très forte contribution au total sont bien sûr inchangées si elles correspondent à des nombres qui sont maintenus tels quels (4.2), sans être groupés.

En revanche, l'addition des effectifs de 9 et 10 permet d'éviter l'apparition de toute contribution élevée dans cette nouvelle colonne 9 + 10.

4.6. Afin de faire apparaître le plus clairement possible les conclusions que nous amène à formuler l'étude du tableau 2, adoptons les conventions suivantes en établissant le tableau n°5. Dans la 1e partie, qui correspond aux cinq groupes d'oeuvres analysées jusqu'à présent, on lira + ou - lorsque z de la case correspondante dépasse $\pm$ 1,96 et reflète dès lors un surplus ou un manque du type métrique considéré. Cette façon permet de montrer tout de suite au lecteur les cases significatives et lui évite la lecture fastidieuse de tous les effectifs de z. Dans les cases où z n'est pas significatif, les mêmes signes + ou - sont encadrés de parenthèses: la tendance à l'abondance ou au déficit est ainsi illustrée.

4.7. Immédiatement, l'on constate une nette opposition entre les deux premiers extraits. Certes, dans les cases 4, 5, 6, 7 puis 11 et 14 n'apparaît aucun signe significatif mais dans les premières au moins les écarts de pourcentages sont, il est vrai, réduits (tableau n°4). L'essentiel est que partout où apparaît dans les deux premières lignes un + ou un - significatif, le signe soit opposé d'une ligne à l'autre.

Les *Epîtres* tiennent un juste milieu. Les différences ne sont-elles pas significatives dans quatre cases seulement? Et ce malgré l'importance de l'oeuvre: avec 3870 mots, elle vient au 2e rang (cf. tableau n°1).

Les *Epigrammes*, par contre, procèdent d'une technique plus négligée: deux termes brefs sont significativement abondants en défaveur de trois types longs. Cette allure se confirme de façon éclatante même dans les cases non significatives: tous les signes sont (-) à partir de la 4e colonne à une seule exception près.

Enfin les *Villes* présentent des écarts significatifs dans des cases opposées. De même que les *Epigrammes* (ligne 4 des tableaux) ressemblaient quelque peu aux oeuvres négligées (ligne 1), de même les *Villes* (ligne 5) rappellent quelque peu les oeuvres soignées (ligne 2) par la disposition des signes dépourvus de parenthèses.

4.8. Notons encore, au terme de ce premier examen, que les différences les plus marquées se manifestent dans deux blocs de colonnes: 1 à 3 et 8 à 14. Entre ces deux ensembles, les colonnes 4 à 7 constituent une sorte de terrain neutre ne comportant dans cette 1e partie du tableau n%5 qu'une seule case significative!

5. *Observation du tableau n°3.*

Cette fois, le *Technopaegnion* est également pris en considération.

5.1. La valeur du χ^2 est, voyons-nous, 772,48 - infiniment significative puisque le χ^2 critique, correspondant à P = 0,05, est 84,5.[8]

5.2. Une remarque s'impose immédiatement: le *Technopaegnion* fournit à lui seul plus de la moitié du χ^2 total.

Ce phénomène se manifeste d'ailleurs régulièrement dans divers types de mots: si le trochée fournit, dans sa colonne 4, une contribution remarquablement élevée, on ne doit pas s'étonner devant les chiffres importants du monosyllabe long (col. 2) car il s'agit là d'une caractéristique singulière de cette oeuvre, sur laquelle l'attention du lecteur a déjà été attirée (3.2). Pareillement, l'amphibraque et le palimbacchée rejoignent le bacchée et le spondée par leur haute contribution. Parmi tous ces types, les monosyllabes longs, amphibraques et palimbacchées s'étaient déjà signalés (4.2) par leur chiffre χ^2 élevé dans le tableau n°2: c'est donc dans l'usage de ces trois schémas métriques que l'on observe le plus aisément d'amples variations.

5.3. La répartition en % (tableau 4) nous renseigne immédiatement sur le sens général des variations: dans le *Technopaegnion*, les types 1 à 5 présentent tous un surplus, parfois écrasant, presque rejoints par l'anapeste (type 7). Inversément, tous les mots de quatre mores et plus (à partir du type 6) sont plus ou moins rares. L'anapeste se distingue dans ce domaine, avec un écart de près d'1/3 % seulement, alors que toutes les autres différences dépassent 1 %: c'est que l'anapeste apparaît plus que de raison dans les fins de vers (cf. § 3.2).

La plus grande parcimonie se fait jour chez les types 9 (⏑ - ⏑ : 0,38 %) et 10 (⏑ - - : 0 %!). C'est que la place de prédilection de ces deux derniers est la finale et que, précisément, Ausone les en exclut, dans le *Technopaegnion*, au profit du monosyllabe (cf. § 3.2 encore).

5.4. Le *Technopaegnion*[9] ne se singularise donc pas uniquement par une profusion de monosyllabes mais bien par un surplus de tous les mots brefs, compensé par un manque de tous les types longs - surtout à partir du type 8: le tableau n°5, 2e partie, nous l'indique très clairement. De surcroît, les écarts sont très grands. J'indiquais précédemment (§ 4.6) que j'évitais au lecteur la liste fastidieuse de tous les effectifs de z.

Que l'on me permette toutefois, pour illustrer les fortes variations, de les citer ici à titre d'exemple:

type	z	type	z
1	+ 3,29	8	- 5,24
2	+ 8,63	9	- 6,44
3	(+ 1,15)	10	- 7,37
4	+ 16,53	11	- 5,31
5	(+ 1,70)	12	- 4,44
6	+ 7,34	13	- 2,39
7	(- 0,55)	14	- 3,64

L'opposition entre mots brefs et longs est on ne peut plus nette.

Si l'on effectue un test supplémentaire en regroupant les cinq premières oeuvres pour les opposer aux *Technopaegnion*, on obtient évidemment les mêmes valeurs de z, de signe opposé deux à deux, ce qui donne

	1	2	3	4	5	6	7	8 à 14
Techn.	-	-	(-)	-	(-)	+	(+)	chacun +
autres	+	+	(+)	+	(+)	-	(-)	chacun -

L'originalité du *Technopaegnion* en devient éclatante! Elle est à ce point forte que le χ^2 obtenu lors de ce test supplémentaire est de 542,61: pour 13 degrés de liberté (deux lignes horizontales et 14 types de mots - cf. § 1.2), c'est phénoménal.

5.5. Cet afflux de mots brefs, entraînant fatalement une raréfaction des termes plus longs, provoque logiquement un accroissement net du nombre global de mots, ce que nous montrent les calculs suivants:

a) sans *Technopaegnion*:

	nombre de vers n_i	nombre de vers %	nombre de mots observé	nombre de mots prop. % vers	différence
négl.	1001	37,2	5990	5966	+ 24
soig.	582	21,6	3343	3455	-112
Epist.	657	24,4	3870	3903	- 33
Epigr.	282	10,5	1807	1679	+128
Ordo V.	168	6,2	985	992	- 7
	2690	100	15995	15995	0

Les oeuvres soignées, comme on l'a déjà dit (§ 3.1 et surtout 4.4) présentent moins de mots que prévu, puisque les termes longs l'emportent sur les courts, à l'inverse de ce que l'on observe dans les poèmes négligés.

b) avec *Technopaegnion:*

	nombre de vers n_i	%	nombre de mots observé	prop.% vers	différence
négl.	1001	35	5990	6068	- 78
soig.	582	20,5	3343	3537	-194
Epist.	657	23	3870	3987	-117
Epigr.	282	9,9	1807	1715	+ 92
Vrb.	168	5,9	985	1023	- 38
Techn.	163	5,7	1323	988	+335 !
	2853	100	17318	17318	0

Cette fois, le *Technopaegnion* accapare à lui seul la quasi totalité du surplus. Seules, les *Epigrammes* présentent encore un écart positif, moindre que lors du calcul précédent.

Inversément, c'est dans ses oeuvres "soignées" et, à un degré moindre, dans les *Epitres* ou les *Villes*, qu' Ausone a particulièrement développé l'usage des mots longs: ceux-ci par leur étendue, contribuent à dégager une impression d'ampleur, de majesté, tandis qu'une abondance de termes brefs confèrent aux *Epigrammes* un rythme saccadé, plus proche de la langue parlée.

5.6. Aussi doit-on faire preuve de prudence dans l'emploi des termes "soigné" et "négligé". Ce que l'on appelle "négligence" peut procéder d'une loi du genre - ainsi en est-il du style satirique et épigrammatique.

Par exemple, il serait absurde de qualifier le *Technopaegnion* de poème négligé. Si la statistique décèle une préférence pour les mots brefs et une répugnance - toute relative - aux termes longs, c'est-à-dire une parenté avec ce que nous avons défini comme "style négligé", ce que l'on vient de dire de cette oeuvre nous autorise à conclure que cette répartition résulte d'une volonté délibérée du poète.

Quant à cette séparation entre vers "soignés" et poèmes "négligés", elle ne se manifeste pas uniquement à l'intérieur de l'oeuvre d'Ausone: Prudence en montre également l'exemple, de même que de nombreux auteurs de l'âge classique - les recherches auxquelles je me livre actuellement le prouvent[10].

6. *Regroupements de types métriques équivalents.*

6.1. Les χ^2 totaux des tableaux 2 et 3 sont très élevés. Nous allons essayer de les réduire en regroupant des types de mots interchangeables.

L'expérience a tout d'abord été tentée avec les mots équivalent à deux longues (situés col. 6 à 8 = types - -, ∪∪ -, - ∪∪), mais la différence entre ceux-ci reste significative. Il en est de même avec les mots de trois longues (col. 12, 13 et une partie des effectifs de la col. 14: resp. types - - -, - ∪∪ - et, col. 14: - - ∪∪ + ∪∪ - - + ∪∪ - ∪∪).

6.2. Bien que ces mots soient, nous le répétons, interchangeables d'un point de vue strictement technique, Ausone ne les emploie donc pas indistinctement. En effet, le type 6 comporte deux longues (- -) tandis que les types 7 et 8 comptent une longue et deux brèves (resp. ∪∪ - et - ∪∪) le rythme obtenu est dès lors dissemblable et entraîne des exigences dans le choix du vocabulaire. Ainsi, le test portant uniquement sur les types 6 à 8 du tableau n°2 se révèle significatif à P = 0,02. Mais dès que l'on considère également le *Technopaegnion* et, pour les mots de trois longues, dans tous les cas, l'écart est très significatif.

6.3. Les résultats sont, par contre, plus intéressants lorsque l'on additionne les effectifs des colonnes 9 et 10 - ainsi qu'il a été procédé dans les tableaux 1 à 5.

Les deux types ici considérés se rencontrent, sauf exception, en fin de vers où la quantité de la dernière syllabe est indifférente. De plus, il s'agit souvent de formes variables des mêmes mots: aucun argument rythmique ou linguistique ne permet dès lors d'accroître délibérement les effectifs de l'un ou l'autre type. Aussi est-il tout à fait logique que, dans les tableaux n°2 et 5 (1e partie), le résultat de cette nouvelle colonne soit non significatif et qu'il contribue considérablement à faire baisser le χ^2 total.

Par contre, lorsque l'on incorpore le *Technopaegnion* (cf. tableaux n°3 et 5 (2e partie)), l'écart devient à nouveau manifeste puisque cette oeuvre est très dépourvue des types 9 et 10 (υ - υ et υ - -) à tel point que les *Epitres* deviennent significatives, mais dans le sens inverse bien sûr.

6.4. Le tableau n°6 résume les résultats de ces tentatives de regroupement. Le z de chaque case a tout d'abord été calculé en limitant le tableau à 5 colonnes:

- les mots brefs: monosyllabes brefs et longs[11] et pyrrhiques, ces derniers étant équivalents métriques des monosyllabes longs;
- les mots de 3/4 de pied: trochées et iambes;
- les mots équivalent à 1 pied: spondées, anapestes et dactyles;
- les types 9 et 10 évoqués ci-dessus;
- tous les termes plus longs.

On a ensuite condensé davantage encore pour obtenir trois colonnes: les deux premières restaient inchangées tandis que tous les types restant, 6 à 14, étaient regroupés.

Cette dernière opération m'a semblé utile afin d'éviter la lecture de signe opposés dans les colonnes 6-8, 9-10 et 11-14. Certes, un premier mouvement d'ensemble se manifestait déjà: les cases significatives arboraient le même signe à l'intérieur d'une oeuvre donnée. Ainsi, dans la 1e partie du tableau n°6, localise-t-on les + dans les pièces soignées, puis, dans la 2e partie du tableau, deux + supplémentaires dans les *Epistulae* et trois - dans le *Technopaegnion*.

La contraction en trois colonnes éclaire à nouveau, dans la 1e partie du tableau,

- la similitude entre poèmes négligés et *Epigrammes*,
- le parallélisme relatif entre vers soignés (- et + aux extrêmes) et les *Villes* (- et (+) aux mêmes postes),

- le juste milieu, la position indéterminée des *Epitres*.

Les différences sont dans l'ensemble déjà bien nettes: huit cases sur quinze sont significatives.

Lorsque nous faisons intervenir le *Technopaegnion*, dans la 2e partie, la situation évolue bien sûr: 14 cases sur 18 reçoivent un z supérieur en valeur absolue à 1,96. Dans les oeuvres soignées, Ausone montre une nette aversion pour les termes brefs (z = -7,57); cet ensemble de poèmes est d'ailleurs le seul où le poète confère la primauté au spondée et non au monosyllabe long qu'il préfère dans toutes les autres oeuvres (cf. tableau n°4). Inversément, l'auteur y recherche les mots longs (z = + 9,26), en particulier les types 11 à 14, puis-je préciser, car avant la réduction en un tableau de 3 colonnes, les z étaient de + 4,74 pour les types 6 à 8, + 1,17 pour 9 et 10 et déjà + 5,67 pour 11 à 14.

Inversément, les *Epigrammes* sont garnies de mots brefs (z = + 5,02) et le *Technopaegnion* crève tous les plafonds avec des z de + 9,43, + 15,23 et - 20,02. Avant la réduction en 3 colonnes, ce dernier chiffre était remplacé par - 9,14 pour les types 6 à 8, -9,96 pour 9 et 10, - 8,50 pour 11 à 14. Et je me limite à citer ici les z supérieurs en valeur absolue à 5 tant les différences observées sont fortes!

7. *Conclusion*

S'il en était encore besoin, voici mis en évidence tout au long de ces pages le danger que fait courir l'appréciation globale d'un résultat: idéalement il faut examiner chaque tableau - quelle qu'en soit l'étendue - case par case, avec la prudence nécessaire. Et de surcroît il faut tenir compte du choix des oeuvres à l'examen desquelles on procède. L'exemple du *Technopaegnion* le prouve à souhait: l'insertion de ce poème dans le tableau n°3 modifie grandement les résultats obtenus au tableau n°2; certes, son caractère excentrique n'en ressort que plus facilement mais, comme je viens de le dire, illustre combien il eût été fallacieux de porter un jugement d'ensemble sur l'oeuvre d'Ausone sans opérer en son sein les distinctions opportunes.

Les lumières de la statistique sont en effet indispensables pour éclairer les types métriques que le poète favorise ou au contraire évite de façon significative dans ses différentes oeuvres. Le bas du tableau n°4 nous montre les confusions auxquelles ne manque pas de mener une extrapolation exagérée, basée sur la seule comparaison des fréquences.

Les trois types de mots les plus souvent représentés sont les mêmes dans toutes les oeuvres, dans un ordre 2 - 6 - 4 qui n'enregistre de permutation que dans les poèmes soignés et le *Technopaegnion*. Puis les anapestes et les dactyles se succèdent tour à tour, suivis d'un 3e groupe où figurent pêle-mêle les col. 3, 5, 12 et 14 tandis que le peloton des types 1, 9, 10, 11 et 13 ferme la marche. Des cadres signalent les types qui franchissent quelque peu la limite séparant les deux derniers groupes, par deux fois seulement.

Si nous considérons enfin le *Technopaegnion*, nous ne serons pas surpris d'y constater le grand désordre: ce n'est qu'ici qu'Ausone relègue le spondée à la 3e place, qu'il propulse le pyrrhique au 5e rang, qu'il rejette le dactyle au 7e et les mots longs au 10e tandis que les monosyllabes brefs grimpent au 8e poste. Quelle permutation dans l'ordre habituel!

Cette mise au point étant faite, la présente étude permettra, j'espère, d'éclairer le lecteur sur les deux points que l'on se promettait d'examiner (§ 1):

1°) La comparaison des différentes oeuvres d'Ausone nous amène à la conclusion que les oeuvres "négligées" abondent en mots brefs, et accusent un déficit de termes longs, et que les poèmes "soignés" présentent le phénomène inverse. Parmi ces derniers, on compte la *Moselle* et le *Cupidon* essentiellement, dont semblent se rapprocher les *Villes*. Au nombre de celles-là, en revanche, figurent divers extraits groupés sous le lemme 'négligés', les *Epigrammes* et, moyennant les restrictions développées précédemment (§ 5.7), le *Technopaegnion*.

2°) Le regroupement de mots métriquement équivalents permet de conserver un écart significatif au sein même de ces regroupements, si les types considérés sont sémantiquement et rythmiquement différents - tels les mots de deux et trois *morae*.

Par contre, la fusion des types 9 et 10 mène à une baisse du χ^2 global: c'est que ces schémas, parfaitement interchangeables, sont employés quasi indifféremment par Ausone.

L'aspect positif de ces conclusions se révèle encourageant: il ne manque plus que de procéder à une enquête similaire, tant parmi d'autres écrivains latins - poètes ou prosateurs - que parmi des témoins d'autres littératures. Les dépouillements déjà publiés ainsi que les recherches en cours, font attendre avec impatience des résultats que l'on espère prometteurs.

Notes

[1] Soit les livres 2 à 5 et 7 de l'édition de R. PEIPER dans la collection Teubner (Leipzig 1886), repris tout récemment par S. PRETE (même collection, 1978) sous les numéros 2, 8 à 12, 3, 4 et 13.

[2] En fait, la qualification de ces deux groupements repose sur des critères objectifs: cf. par ex. TORDEUR (1970: 968-969). Cette répartition a déjà été effectuée précédemment lors du choix des textes qui servent de base à ma thèse de doctorat: cf. par ex. Computer and the Humanities 5 (1970-71) fasc. 2, p.103, réf. L. 389.

[3] Le titre n'en signifie-t-il pas "jeu d'art", "excercice de technique (poétique)"?

[4] En outre, le *Cento Nuptialis*, fait d'hémistiches virgiliens, ne peut de ce fait véritablement figurer parmi les oeuvres d'Ausone.

[5] Pour ces données théoriques ainsi que pour la réalisation pratique des calculs, j'exprime avec reconnaissance ma dette envers M. Edmond ROBAYE (Université Libre de Bruxelles) qui a bien voulu consacrer son attention à leur mise au point. J'ai déjà tenté une première expérience (la méthode choisie était beaucoup plus simple) sur les cinq grands poèmes en hexamètre de Prudence: cf. TORDEUR (1972).

Tout bon manuel de statistique explique et illustre le maniement de ces formules; en langue française, on peut conseiller particulièrement l'ouvrage de Ch. MULLER (1968), particulièrement les pages 98 à 100 et la table p.239, repris et amplifié en trois volumes: MULLER (1973, 1978a, 1978b). Dans la dernière formule citée en ces pages, n représente le nombre de "degrés de libertés", c'est-à-dire (h-1) x (l-1), h représentant le nombre de cases du tableau en "hauteur" et l le nombre correspondant en "largeur". Dans notre tableau n°1 par ex., avant addition du *Technopaegnion*, on analyse 14 types en cinq oeuvres; on en déduit donc n = (14-1) x (5-1) ou 13 x 4 = 52. Les tables donnent z correspondant à P = 0,05 la valeur de 1,645. Toutes ces données permettent alors le calcul du χ^2. Pour tester ensuite chaque case, on a alors suivi la méthode exposée par RAO (1973: 393).

6 Seul ∪∪ - - se rencontre aussi occasionellement en fin de vers. On emploiera dans la présente étude les schémas traditionnels ∪ = brève, - = longue. En outre, ∪∪ = -.

7 cf. TORDEUR (1972) § 5.2 et note 5. Pour donner une idée plus précise du degré de signification atteint ici: de la formule citée plus haut (§ 1.2) on tire z

$$z = \sqrt{2\chi^2} - \sqrt{2n-1} = 11,18.$$

On sait que z = ± 1,96 correspond à P = 0,05, le seuil de signification que nous nous sommes fixé. P devient très rapidement minime:

z = ± 2,326 indique P = 0,02
z = ± 2,576 indique P = 0,01
z = ± 3,291 indique P = 0,001
z = ± 3,891 indique P = 0,0001
z = ± 4,417 indique P = 0,00001.

Dès que z atteint 4, P est donc déjà inférieur à 10^{-4}.

8 C'est le moment de rappeler que *Technopaegnion* signifie "exercice de technique (poétique)" (cf. § 1.1 in fine et note 3) - pour autant, bien sûr, que l'on puisse proposer une traduction assurée pour ce terme.

9 cf. note 2. En outre, la répartition des différents types de mots ne constitue bien entendu qu'un critère parmi d'autres: dans le domaine de la métrique, des écarts surviennent également dans la fréquence, la place et le type des élisions, les arrangements verbaux en fin de vers, ... de même que dans le choix du vocabulaire - par exemple. Mais ces considérations, pour intéressantes qu'elles soient, débordent du cadre de la présente étude.

10 Il est intéressant de constater que le poète Prudence adopte une attitude quasiment semblable, même si la méthode pour l'apprécier est plus simple (cf. TORDEUR 1972: § 6.7 et 6.8). Les autres écrivains auxquels je travaille actuellement se comportent souvent de la même façon, mais il est prématuré de tirer dès à présent des conclusions d'ensemble.

11 Certes, le tableau des fréquences (tableau n°4, bas) indique la grande différence de comportement entre ces deux types métriques: ils sont presque diamétralement opposés. Je les regroupe ici pour des raisons morphologiques: ce sont bien les mêmes mots que les circonstances de la prosodie rendent tantôt brefs tantôt longs. Le fait que la colonne 1-3 soit presque toujours significative montre que ce regroupement n'a pas pour effet d'atténuer les différences, tant est forte la prédominance du monosyllabe long.

Bibliographie

MULLER, C., Initiation à la statistique linguistique. Paris: Larousse 1968

------, Initiation aux méthodes de la statistique linguistique. Paris: Hachette 1973

------, Principes et méthodes de statistique lexicale. Paris: Hachette 1978a

------, Statistique linguistique. Exercices et applications. Paris: Hachette 1978b

RAO, C.R., Linear Statistical Inference and Its Applications. New York: Wiley 1973

TORDEUR, P., Etude des élisions dans la Moselle d'Ausone. Latomus 29, 1970, 966-987

------, Essai d'analyse statistique de la métrique de Prudence, Organisation internationale pour l'étude des langues anciennes par ordinateur - Revue 1972 fasc. 2, 19-37

Tableau 1. Ausone : Effectifs observés : effectifs bruts

types prosodiques	1	2	3	4	5	6	7	8	9	10	11	12	13	14		9+10
	⏑	–	⏑⏑	–⏑	⏑–	––	⏑⏑–	–⏑⏑	⏑–⏑	⏑––	––⏑	–––	–⏑⏑–	⏑–⏑⏑ ––⏑⏑ ⏑⏑–– ⏑⏑–⏑⏑	TOTAL	⏑–⏑ ⏑––
négligé	122	1244	361	630	290	938	439	521	244	206	206	346	163	280	5990	450
soigné	102	512	151	346	160	548	268	330	90	165	142	224	125	180	3343	255
Epist.	95	732	206	421	236	575	292	288	157	168	181	190	136	193	3870	325
Epigr.	53	422	125	191	76	275	128	176	67	58	38	95	34	69	1807	125
Ordo V.	38	155	40	124	46	143	76	90	45	35	42	54	44	53	985	80
TOTAL	410	3065	883	1712	808	2479	1203	1405	603	632	609	909	502	775	15995	1235
Techn.	54	384	83	344	81	106	94	61	5	0	13	37	26	35	1323	5
TOTAL GEN.	464	3449	966	2056	889	2585	1297	1466	608	632	622	946	528	810	17318	1240

Tableau 2. Ausone : Chi carré (*Techn.* exclus)

types	1	2	3	4	5	6	7	8	9	10	11	12	13	14	TOTAL	9+10	TOTAL
négligé	6,48	8,06	2,78	0,19	0,52	0,10	0,29	0,05	1,46	3,98	2,13	0,09	3,32	0,36	29,83	0,34	24,73
soigné	3,10	25,81	6,10	0,39	0,47	1,72	1,09	4,50	10,30	8,20	1,70	6,09	3,84	2,01	75,33	0,04	56,87
Epist.	0,18	0,12	0,27	0,11	8,39	1,03	0,00	7,94	0,85	1,49	7,69	4,07	1,74	0,16	34,04	2,30	34,00
Epigr.	0,96	16,57	6,39	0,03	2,56	0,09	0,46	1,88	0,02	2,51	13,79	0,58	9,10	3,93	58,86	1,51	57,84
Ordo V.	6,44	6,03	3,80	3,27	0,28	0,61	0,05	0,14	1,67	0,39	0,54	0,07	5,54	0,58	29,42	0,20	27,57
	17,16	56,60	19,34	4,00	12,22	3,55	1,90	14,51	14,29	16,57	25,85	10,90	23,54	7,04	227,49	4,39	201,01

Tableau 3. Ausone : Chi carré (*Techn.* inclus)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	TOTAL	9+10	TOTAL
négligé	9,23	2,18	2,16	9,26	0,99	2,15	0,21	0,38	5,40	0,73	0,39	1,08	2,11	0,00	36,28	1,04	31,19
soigné	1,73	35,52	6,75	6,52	0,79	4,81	1,24	7,81	6,38	15,16	4,01	9,38	5,22	3,57	108,89	1,02	88,37
Epist.	0,73	1,95	0,45	3,22	7,02	0,01	0,02	4,79	3,29	5,07	12,69	2,17	2,75	0,79	44,94	8,28	44,86
Epigr.	0,43	10,72	5,81	2,58	3,03	0,10	0,40	3,47	0,20	0,96	11,15	0,14	8,08	2,85	49,92	0,15	48,91
Ordo V.	5,11	8,64	4,06	0,43	0,41	0,11	0,07	0,53	3,14	0,02	1,24	0,00	6,50	1,04	31,30	1,27	29,40
Techn.	9,71	55,12	1,15	222,48	2,52	42,38	0,26	23,22	36,99	48,28	25,07	17,21	5,10	11,68	501,16	84,99	500,89
	26,94	114,14	20,39	244,48	14,76	49,57	2,19	40,19	55,39	70,22	54,55	29,98	29,75	19,94	772,48	96,75	743,62

TABLEAU 4 A. Conversion en pourcentage par oeuvre (cf. tableau 1)

types	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Σ	9+10
négligé	2,04	20,77	6,03	10,52	4,84	15,66	7,33	8,70	4,07	3,44	3,44	5,78	2,72	4,67	100	7,51
soigné	3,05	15,32	4,52	10,35	4,79	16,39	8,02	9,87	2,69	4,94	4,25	6,70	3,74	5,38	100	7,63
Epist.	2,45	18,91	5,32	10,88	7,00	14,86	7,54	7,44	4,06	4,34	4,68	4,91	3,51	4,99	100	8,40
Epigr.	2,93	23,35	6,92	10,57	4,21	15,22	7,08	9,74	3,71	3,21	2,10	5,26	1,88	3,82	100	6,92
Ordo V.	3,86	15,74	4,06	12,59	4,67	14,52	7,72	9,14	4,57	3,55	4,27	5,48	4,47	5,38	100	8,12
TOTAL	2,56	19,16	5,52	10,73	5,05	15,49	7,52	8,78	3,77	3,95	3,81	5,68	3,14	4,84	100	7,72
Techn.	4,08	29,02	6,27	26,00	6,12	8,01	7,11	4,61	0,38	0,00	0,98	2,80	1,96	2,64	100	0,38
TOTAL GEN.	2,68	19,91	5,58	11,87	5,13	14,92	7,49	8,46	3,50	3,65	3,59	5,46	3,05	4,68	100	7,15

Tableau 4 B. Ordre décroissant de fréquence des schémas métriques

Rang	1	2	3	4	5	6	7	8	9	10	11	12	13	14
négligé	2	6	4	8	7	3	12	5	14	9	10	11	13	1
soigné	6	2	4	8	7	12	14	10	5	3	11	13	1	9
Epist.	2	6	4	7	8	5	3	14	12	11	10	9	13	1
Epigr.	2	6	4	8	7	3	12	5	14	9	10	1	11	13
Ordo V.	2	6	4	8	7	12	14	5	9	13	11	3	1	10
TOTAL	2	6	4	8	7	12	3	5	14	10	11	9	13	1
Techn.	2	4	6	7	3	5	8	1	12	14	13	11	9	10
TOTAL GEN.	2	6	4	8	7	3	12	5	14	10	11	9	13	1

Tableau 5. Cases significatives du tableau 2

1e partie: *Techn.* exclus

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	9+10
nég-ligé	-	+	+	(-)	(-)	(+)	(-)	(-)	(+)	-	(-)	(+)	-	(-)	(-)
soig-né	+	-	-	(-)	(-)	(+)	(+)	+	-	+	(+)	+	+	(+)	(-)
Epist.	(-)	(-)	(-)	(+)	+	(-)	(+)	-	(+)	(+)	+	-	(+)	(+)	(+)
Epigr.	(+)	+	+	(-)	(-)	(-)	(-)	(+)	(-)	(-)	-	(-)	-	-	(-)
Ordo V.	+	-	-	(+)	(-)	(-)	(+)	(+)	(+)	(-)	(+)	(-)	+	(+)	(+)

2e partie: *Techn.* inclus

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	9+10
nég-ligé	-	+	(+)	-	(-)	+	(-)	(+)	+	(-)	(-)	(+)	(-)	(-)	(+)
soig-né	(+)	-	-	-	(-)	+	(+)	+	+	+	+	+	+	+	(+)
Epist.	(-)	(-)	(-)	-	+	(-)	(+)	-	+	+	+	(-)	(+)	(+)	+
Epigr.	(+)	+	+	(-)	(-)	(+)	(-)	+	(+)	(-)	-	(-)	-	(-)	(-)
Ordo V.	+	-	-	(+)	(-)	(-)	(+)	(+)	(+)	(-)	(+)	(+)	+	(+)	(+)
Techn.	+	+	(+)	+	(+)	-	(-)	-	-	-	-	-	-	-	-

Tableau 6 : Regroupement successifs de types de mots

1e partie: *Techn.* exclus

	1 - 3	4 - 5	6 - 8	9-10	11-14	1 - 3	4 - 5	6-14
négligé	+	(-)	(-)	(-)	-	+	(-)	-
soigné	-	(-)	+	(-)	+	-	(-)	+
Epist.	(-)	+	-	(+)	(+)	(-)	+	(-)
Epigr.	+	(-)	(+)	(-)	-	+	(-)	-
Ordo V.	-	(+)	(-)	(+)	(+)	-	(+)	(+)

2e partie: *Techn.* inclus

	1 - 3	4 - 5	6 - 8	9-10	11-14	1 - 3	4 - 5	6-14
négligé	(+)	-	(+)	(+)	(-)	(+)	-	(+)
soigné	-	-	+	(+)	+	-	-	+
Epist.	-	(-)	(-)	+	+	-	(-)	+
Epigr.	+	-	(+)	(-)	-	+	-	-
Ordo V.	-	(+)	(+)	(+)	+	-	(+)	+
Techn.	+	+	-	-	-	+	+	-

A NOTE ON POL TORDEUR'S PAPER "ETUDE STATISTIQUE SUR L'HEXAMÈTRE D'AUSONE"

Rüdiger Grotjahn, Bochum

In the following note on P. TORDEUR's paper "Etude statistique sur l'hexamètre d'Ausone" (pp. 75-96) in this volume) I first want to comment on the test for deviation in a single cell of a contingency table used by TORDEUR in his attempt to determine the style of several poems of Ausonius according to the frequency of different metrical word types. I then want to show that, without the computation of appropriate coefficients which measure the relationship between two or more variables independently from the sampling error, in the analysis of contingency tables one may arrive at conclusions which may be not completely justified.

1. Test for Deviation in a Single Cell

In extracting information from his Table 1 (p. 92 in this volume), TORDEUR interprets not only the contribution of each cell to the X^2 value obtained for the whole table with the help of the statistic $(o_i-e_i)^2/e_i$, but he also tests the deviation observed for significance (cf. TORDEUR's Table 5). Since neither TORDEUR himself nor most statistical textbooks explain the test for deviation in a single cell, a short comment on this method seems to be indicated. (Readers interested in more detailed information are referred to HABERMAN (1973) and ŘEHÁK/ŘEHÁKOVÁ (1980).)

Let us illustrate the use of the test for deviation in a single cell by means of the second cell of the row entitled "*Egigr.*" in TORDEUR's Table 1. In order to test whether the observed frequency of 422 deviates from the value expected under the null hypothesis of independence between the variables 'poem' and 'metrical word type', we pool those frequencies which are the result of a cross-classification of the metrical word types 1,3,4,..,14 and the texts "négligé", "soigné", "*Epist.*" "*Ordo V.*", and "*Techn.*". In this way we obtain from TORDEUR's Table 1 the following 2x2 table:

Table 1. 2x2 Table Obtained from TORDEUR's Table 1

poem	metrical word type 2	others	TOTAL
Epigr.	422	1385	1807
others	3027	12484	15511
GENERAL TOTAL	3449	13869	17318

This 2x2 table can now be tested for significance in several different ways. We can, for example, use the Fisher-Yates test, which is an exact procedure, or the various approximative methods such as the χ^2 test or the z test (cf. the discussion of these procedures in R. GROTJAHN, "A Statistical Model for the Coincidence of Ictus and Accent", pp. 33-74 in this volume).

Since TORDEUR uses the z test I will confine myself to explaining this procedure only. Its application to the above 2x2 table as well as to the remaining 73 cells of TORDEUR's Table 1 is justified because the frequencies are sufficiently large. The z test can be very easily calculated, and compared to the χ^2 test has the advantage that the sign of z indicates the direction of the difference between observed and expected frequencies.

There are several equivalent formulas which can be used in the application of the z test. We will present only two, both of which are very simple to calculate on the basis of contingency tables such as TORDEUR's Table 1.

The two formulas can be derived from the well known χ^2 statistic for 2x2 tables

$$\chi^2 = n \frac{(ad - bc)^2}{(a+b)(a+c)(c+d)(b+d)}, \qquad (1)$$

if we denote the frequencies of a 2x2 table not by a, b, c, and d, but as it is shown in Table 2.

Table 2. General Form of a 2x2 Table

n_{ij}	$n_{i.}-n_{ij}$	$n_{i.}$
$n_{.j}-n_{ij}$	$n-n_{i.}-n_{.j}+n_{ij}$	$n-n_{i.}$
$n_{.j}$	$n-n_{.j}$	n

Since, for df=1, $z=\sqrt{\chi^2}$, it follows from (1) that

$$z = \sqrt{n} \frac{ad - bc}{\sqrt{(a+b)(a+c)(c+d)(b+d)}},$$

which after some simple rearrangements can be written using the notation of Table 2 as

$$z_{ij} = \sqrt{n} \frac{nn_{ij} - n_{i.}n_{.j}}{\sqrt{n_{i.}n_{.j}(n-n_{.j})(n-n_{i.})}} \qquad (2a)$$

or as

$$z_{ij} = n \frac{n_{ij} - \frac{n_{i.}n_{.j}}{n}}{\sqrt{\frac{n_{i.}n_{.j}}{n}(n-n_{.j})(n-n_{i.})}}. \qquad (2b)$$

Applying (2a) to the data of Table 1 we obtain

$$z = \sqrt{17318}\, \frac{17318(422) - 1807(3449)}{\sqrt{1807(3449)\ 13869(15511)}} = 3.87.$$

Let us now illustrate how (2) can be used if we do not previously establish a 2x2 table. Using (2b) we obtain e.g. for the metrical word type 6 in the row "*Techn.*" of TORDEUR's Table 1

$$z = 17318 \frac{106 - \frac{1323(2585)}{17318}}{\sqrt{\frac{1323(2585)}{17318}14733(15995)}} = -7.3,$$

which is exactly the same value as given by TORDEUR in his first table on page 83.

If we compute the z statistic for all cells in TORDEUR's Table 1, we get the data displayed in Table 3.

Table 3. Z Scores Obtained in the Test for Deviation in a Single Cell

types	négligé	soigné	*Epist.*	*Epigr.*	*Ordo V.*	*Techn.*
1	-3.8	1.5	-1.0	0.7	2.4	3.3
2	2.0	-7.4	-1.8	3.9	-3.4	8.6
3	1.9	-3.0	-0.8	2.6	-2.1	1.1
4	-4.0	-3.0	-2.2	-1.8	0.7	16.5
5	-1.3	-1.0	3.1	-1.9	-0.7	1.7
6	2.0	2.6	-0.1	0.4	-0.4	-7.3
7	-0.6	1.3	0.1	-0.7	0.3	-0.6
8	0.8	3.3	-2.6	2.1	0.8	-5.2
9	2.9	-2.9	2.1	0.5	1.9	-6.4
10	-1.1	4.4	2.6	-1.1	-0.2	-7.4
11	-0.8	2.3	4.1	-3.6	1.2	-5.3
12	1.3	3.5	-1.7	-0.4	0.0	-4.4
13	-1.8	2.6	1.9	-3.0	2.7	-2.4
14	-0.0	2.2	1.0	-1.8	1.1	-3.6

In interpreting the results of Table 3 we have to take into account, however, that we have carried out a total of 74 tests. If we consider these tests as a "family" of simultaneous tests (cf. MILLER 1966 and MATTHÄUS 1980), the actual significance level may be considerably larger than the nominal one, which has been set at $\alpha = 0.05$ by TORDEUR. If we require that all 74 tests should have the same significance level the nominal level α may be adjusted by dividing it by the number of simultaneous tests. If we denote this number by r, the adjusted level is then

$$\alpha' = \alpha/r. \tag{3}$$

In the present case we obtain $\alpha' = 0.05/74 = 0.0006$ which corresponds to $|z| = 3.4$, a value which is much larger than $|z| = 1.96$ used by TORDEUR. The effect is that only 18 cells show significant deviations, whereas in TORDEUR's analysis 44 cells

turned out to be significant (cf. TORDEUR's Table 5, 2nd Part). Furthermore, out of the 18 significant cells there are 9 from *Technopaegnion*, which seems to indicate that this poem differs considerably from the other ones.

Moreover, if we compare the number of significant cells in the row "soigné" with that obtained by TORDEUR, we realize that there are only 3 significant cells whereas in TORDEUR's Table 5 there are 11 significant z scores in the corresponding row. As this example shows, the adjustment of α may lead to considerably different results and thus to different interpretations as well.

2. Measuring the Strength of the Relationship between the Variables 'Metrical Word Type' and 'Poem'

In interpreting the z test as well as the various χ^2 tests carried out by TORDEUR we have to take into consideration that the magnitude of the z or X^2 values obtained depends on the sample size.[1] The larger the sample size, the larger the z or X^2 values even if the strength of the relationship between the variables is the same in the tables under investigation. Thus, interpreting highly significant z or X^2 values as indicating, for instance, strong differences between two or more poems may be not correct.

There exist, however, various contingency coefficients which allow us to measure the strength of the relationship between two or more variables independently from the sample size. One of these measures is the coefficient C_c introduced by CRAMÉR which is very practical if we have already calculated X^2 values. C_c is defined (cf. e.g. BLALOCK 1979: 305) as

$$C_c = \sqrt{\frac{X^2}{n\cdot\min(r-1,c-1)}}\,, \tag{4}$$

where r designates the number of rows and c the number of columns in a rxc contingency table. C_c can vary from zero to unity and becomes equivalent to the well known φ coefficient in the case a of 2x2 table. C_c is zero if there is absolute agreement between observed and expected frequencies and approaches unity if the deviations increase.

Let us now calculate some C_c coefficients for TORDEUR's data. Taking the *Technopaegnion* into account TORDEUR has obtained for his Table 1 a X^2 value of 772.48, and excluding the *Technopaegnion* a value of 227.49 (cf. TORDEUR's Tables 2 and 3). From these two X^2 values we obtain for the table including *Technopaegnion*

$$C_c = \sqrt{\frac{772.48}{17318(5)}} = 0.0945$$

and for the table without *Technopaegnion*

$$C_c = \sqrt{\frac{227,49}{15995(4)}} = 0.0596.$$

Although both values are extremely significant ($P \ll 0.1 \cdot 10^{-8}$ in both cases) both coefficients indicate a very high degree of homogeneity in TORDEUR's Table 1. Since the value obtained without taking *Technopaegnion* into account is only half of that obtained in the analysis including *Technopaegnion*, this poem seems to differ to some extent from the other ones. We therefore compute C_c from the extremely high X^2 value of 542.61 obtained by TORDEUR in comparing *Technopaegnion* to the pooled remaining poems (cf. TORDEUR p. 83):

$$C_c = \sqrt{\frac{542.461}{17318}} = 0.1770.$$

Although this value is considerably higher than $C_c = 0.0596$ obtained for TORDEUR's Table 1 excluding *Technopaegnion*, TORDEUR's interpretation of the X^2value as indicating a very high degree of originality is in my opinion not justified.

This conclusion is supported if we calculate KENDALL's coefficient of concordance W (cf. KENDALL 1962 or GIBBONS 1971: 250ff)

for the rank ordered data of TORDEUR's Table 4B. W is defined as follows:

$$W = \frac{12}{c^2(r^3-r)} \sum_{i=1}^{r} \left(R_{i.} - \frac{c(r+1)}{2}\right)^2, \qquad (5)$$

where $R_{i.}$ designates the sum of the ranks in column i.

In order to compute W, we first arrange TORDEUR's Table 4B in a slightly different way. We assign the rank 1 to the smallest frequency of a row, the rank 2 to the next smallest and so forth. In this way we obtain the rank orders represented in Table 4.

Table 4. Rank Order of the Relative Frequencies in TORDEUR's Table 4B

types	negligé	soigné	*Espist.*	*Epigr.*	*Ordo V.*	$R_{i.}$	*Techn.*	$R_{i.}$
1	14	13	14	12	13	66	8	74
2	1	2	1	1	1	6	1	7
3	6	10	7	6	12	41	5	46
4	3	3	3	3	3	15	2	17
5	8	9	6	8	8	39	6	45
6	2	1	2	2	2	9	3	12
7	5	5	4	5	5	24	4	28
8	4	4	5	4	4	21	7	28
9	10	14	12	10	9	55	13	68
10	11	8	11	11	14	55	14	69
11	12	11	10	13	11	57	12	69
12	7	6	9	7	6	35	9	44
13	13	12	13	14	10	62	11	73
14	9	7	8	9	7	40	10	50

We now compute W and obtain for the data including *Technopaegnion*

$$W = \frac{12}{6^2(14^3-14)} \sum_{i=1}^{14} \left(R_{i.} - \frac{6(15)}{2}\right)^2 = 0.8825$$

and for the data without *Technopaegnion*

$$W = \frac{12}{5^2(14^3-14)} \sum_{i=1}^{14} (R_{i.} - \frac{5(15)}{2})^2 = 0.9138.$$

Since for large c and/or r

$$\chi^2 = c(r-1)W \quad \text{with df} = r-1, \tag{6}$$

W can be tested by using χ^2 tables. We obtain $\chi^2 = 6(13)0.8825 = 68.84$ ($P << 0.001$) and $X^2 = 5(13)0.9138 = 59.40$ ($P << 0.001$). Thus, in both cases there is a high degree of concordance (with a very high probability not due to chance) among the different kind of texts with regard to the rank orders of the frequencies of the metrical word types. This conclusion is sustained if we calculate Spearman's rank correlation coefficient r_s for the rank order obtained for *Technopaegnion* and that which we can assign to the row "TOTAL" in TORDEUR's Table 4A. We get $r_s = 0.8242$ which indicates a high degree of agreement between the two rank orders. Since r_s is significant with $\alpha < 0.001$ the difference between the rank order of *Technopaignion* and that of the remaining poems can be attributed with good reason to chance. As is however shown by the highly significant $C_c = 0.1770$ there are nevertheless significant differences in the frequencies between *Technopaegnion* and the other poems, although these differences are not as important as argued by TORDEUR.

Finally I want to examine two other conclusions at which TORDEUR arrives. According to TORDEUR there is a clear difference between the two kinds of text "entitled "négligé" and "soigné" respectively (cf. p. 81). This seems to be correct if we compare the two kinds of texts only with the help of a χ^2 test. We obtain $X^2_{13} = 97.08$, which is highly significant. If we compute C_c however, we get $C_c = \sqrt{97.08/9333} = 0.1020$ - a value which does not support TORDEUR's statement very well.

The second conclusion concerns the metrical word types 9 and 10, which according to TORDEUR are completely interchangeable and show almost no differences in their distributions (cf. TORDEUR p. 89). Comparing these two metrical word types with the help of

a χ^2 test, however, yields $X^2_5 = 32.09$, which is highly significant. From $X^2 = 32.09$ we get $C_c = 0.1609$ - a value which is almost as high as that obtained by comparing *Technopaegnion* with the other poems and higher than that obtained in comparing the rows "négligé" and "soigné". The conclusion that there is a clear difference in the case of *Technopaegnion* and that in the case of the metrical word types 9 and 10 there is almost no difference seems, therefore, to be hardly justified. In my opinion one should rather conclude that in both cases there are significant but not very important differences between the poems. Taking the results of the test for deviation in a single cell into account, the differences between *Technopaegnion* and the other poems seem to be, however, more important than those between the metrical word types 9 and 10.

Note

1 Since the statistic $(o_i-e_i)^2/e_i$ is only approximately distributed as χ^2, I use the notation X^2 instead of χ^2 which is the notation used by TORDEUR.

References

BLALOCK, H.M.Jr., Social Statistics. Rev. 2nd. Edition. Tokyo, McGraw-Hill Kogakusha 1979

GIBBONS, J.D., Nonparametric statistical inference. New York, McGraw-Hill 1971

HABERMAN, S.J., The Analysis of Residuals in Cross-Classified Tables. Biometrics 29, 1973, 205-220

KENDALL, M.G., Rank Correlation Methods. New York, Hafner 1962

MATTHÄUS, W., Diskussion zu Řehák & Řeháková, Analyse von Kontingenztafeln. In: MATTHÄUS, W. (Ed.), Glottometrika 3. Bochum, Brockmeyer 1980, 29

MILLER, R.G., Simultaneous statistical inference. New York, McGraw-Hill 1966

ŘEHÁK, J., ŘEHÁKOVÁ, B., Analyse von Kontingenztafeln: Zwei Grundtypen von Aufgaben und das Vorzeichenschema. In: MATTHÄUS, W. (Ed.), Glottometrika 3. Bochum, Brockmeyer 1980, 1-29

Sentence Length in Greek Hexameter Poetry[1]

Dee Lesser Clayman, New York

Of all genres of Greek poetry, the epic is by far the most conservative. With a fixed metrical line of dactylic hexameters, a vocabulary of standard formulas, and traditional contents inherited from Homer and Hesiod, the form of the Greek *epos* appears to change little through the course of time. A sensitive reader, however, would never mistake an Hellenistic epic for an early one, the *Argonautica* of Appollonius for Homer's *Iliad*, nor would he or she mistake Appollonius' imitations of early hexameters with those of his contemporary Callimachus. Even within the rigid framework of the Greek epic the best poets speak in their own voice.

Since the uniqueness of Greek hexameter poems cannot be defined completely in the usual way by content, diction and meter, it is appropriate to turn to quantitative analysis with the hope of finding that poets express their individuality by arranging traditional elements in statistically unique patterns. Sentence length seems like a promising area of inquiry because it is one aspect of style which a hexameter poet is free to alter at will. A poet may acknowledge the limits of the metrical line by punctuating at line end, to form sentences exactly 1 line, 2 lines or n lines in length, or he may ignore the metrical line entirely to compose short or long sentences which begin and/or end in midverse. When a sentence runs over the line end, the line is said to be enjambed.

I use the word "sentence" advisedly since the Greeks did not have a sense of grammatical completeness such as we expect to find in a sentence. For Aristotle and later rhetorical theorists the smallest unit of composition was the colon, which could be independent or grouped with other cola to form a period (Arist. *Rhet.* 3.9.5-7; Demetrius *Eloc.* 1.9-11). Since ancient authors have left no autograph manuscripts a modern editor can only use his own best judgment in determining the degree to which cola and periods are independent. Different editors, then, punctuate the same text differently.

Measuring sentence length in these circumstances would be all but impossible except for the fact that the differences in punctuation between editions are not usually so drastic as to produce significantly different statistics. The constraints of Greek syntax do, in fact, make most punctuation a matter of necessity rather than choice. Secondly, a deep strain of conservatism in text editing tends to make the punctuation of most new editions of Greek texts conform to that of the old—at the very least, a competent editor will not alter the punctuation of an older edition without serious deliberation. Finally, most differences in punctuation between editions are differences in quality not quantity. One editor will use a strong stop were another will use a weak stop. The amount and placement of the punctuation remain the same. If all stops are made equivalent, i.e. if no distinction is made between the strong stop, weak stop and interrogation mark, editorial differences will be kept to a minimum, (WAKE 1957: 334-337; JANSON 1964). This is a policy followed in the study of sentence length in Greek hexameter poetry reported below.

Before turning to the study itself it is well to consider what outcome philologists would expect. The father of modern Homeric studies is Milman PARRY, who made the first quantitative study of enjambement in Greek hexameter poetry in order "to seize more surely the way in which the thought of the poet unfolds from verse to verse,"(PARRY 1929: 200-220). PARRY compared six samples, each of 100 consecutive lines, from Homer's *Iliad* and *Odyssey*, Apollonius' *Argonautica* and Virgil's *Aeneid*. PARRY's sampling procedure has been justifiably criticized,(CLAYMAN / VAN NORTWICK 1978: 87-88), yet his conclusions have been widely influential. Two are relevant to the question of sentence length. First of all, PARRY found that in Homer nearly one half of the verses finish where the sentence ends (48.5% in the *Iliad*, 44.8% in the *Odyssey*, while Apollonius has about four such verses to Homer's five (34.8% in the *Argonautica*). This difference did not seem to him to be

"striking," (PARRY 1929: 254). More impressive was the difference between Homer and Apollonius in the use of verses which hold just one sentence, or in some cases, two sentences. PARRY found that one line sentences occur in the Homeric poems about every fifth or sixth line, which is about twice as often as in the *Argonautica*, (PARRY 1929: 255).

In a later article, including evidence from south Slavic oral poetry, PARRY linked the use of one line sentences to the mechanics of oral composition:

> The easiest formula for the oral poet to handle is that which is both a whole sentence and a whole verse. This is the only formula which is complete in itself both in rhythm and thought. It is only formulas of this kind which can be joined on to one another, and be joined together in any number to make a shorter or longer passage. Formulas of other kinds can fill only a given part of the verse, and they must be preceded and followed by formulas which are different in rhythm and which contain other parts of the sentence. Thus the art of the oral poet is largely that of grouping together whole fixed verses.
> (PARRY 1933: 194-195)

In this way oral composition accounts for the preponderance both of one line sentences and of sentences ending at verse end in the Homeric poems. It would seem to follow from PARRY's argument that poems which were composed orally, or which are older and therefore closer to the oral tradition, would have a greater proportion of shorter sentences, the length of one verse or less, while later hexameter poems would have more enjambed lines and therefore longer sentences composed with less attention to verse end. PARRY's statistics for the *Argonautica*, which was written in the third century B.C. and is certainly not an oral composition, support this thesis.

In order to test Parry's hypotheses and to learn more about sentence length in Greek hexameter poetry statistics on sentence length were gathered from all Greek hexameter poetry from the eighth to the second centuries B.C. including in their entirety: the *Iliad* and *Odyssey* of Homer, the *Theogony* and *Works and Days* of Hesiod, the *Shield of Heracles*, the four long

Homeric Hymns (the *Hymn to Demeter*, *Hymn to Apollo*, *Hymn to Hermes* and *Hymn to Aphrodite*), the *Hymns* of Callimachus, the *Argonautica* of Apollonius of Rhodes, the *Phaenomena* of Aratus, the *Idylls* of Theocritus and the *Therica* and *Alexipharmaca* of Nicander of Colophon, a total of 23,049 sentences.[2] The poems can be grouped by subgenre and date. Three subgenres are defined by subject matter: the four *Homeric Hymns* and the *Hymns* of Callimachus form the first and most distinctive subgenre; the heroic epics including the *Iliad*, *Odyssey*, *Shield of Heracles* and *Argonautica* comprise a second; and the third contains the *Theogony*, *Works and Days*, *Phaenomena*, *Theriaca* and *Alexipharmaca* which are all, broadly speaking, didactic poems. Theocritus' *Idylls* are *sui generis*. There are two broad chronological groups: the early hexameters (eighth-fifth centuries B.C.) including the *Iliad*, *Odyssey*, *Shield of Heracles*, *Theogony*, *Works and Days* and the four *Homeric Hymns*; and the Hellenistic hexameters (third-second centuries B.C.) including the *Argonautica*, Callimachus' *Hymns*, *Phaenomena*, *Theriaca*, *Alexipharmaca*, and the *Idylls*.

In previous sentence length studies sentences have been measured by counting the number of words each contains (YULE 1938; WILLIAMS 1939; WAKE 1957; MORTON / LEVISON 1966). This method results in the loss of certain important information. In Greek hexameter poetry words are as short as δ' or as long as προπροκυλινδόμενος (*Il.* 22.221; *Od.* 17.525), the longest word in Greek hexameter poetry. A sentence with very short words is much shorter than a sentence containing the same number of long words. This problem could be solved partially by measuring sentence length by number of syllables, (JANSON 1964: 27 nt. 3), but it is also the case that a syllable can have as many as five phonemes or as few as one (D.H. *de Comp. XV*). The ideal way of calculating sentence length ought to be by number of phonemes per sentence. If the counting is done by machine it is no more difficult to count phonemes than to count syllables or words.

In order to provide data for evaluating the relative merits of counting words, syllables and phonemes, separate sentence length statistics have been produced for each of the three entities. Table 1 contains the data for sentence length measured by number of words. A word is defined as a string of characters surrounded by blanks. Enclitics, pre- and postpositives are counted as separate words since their presence or absence is a matter for the author to decide and therefore a component of style. The works are arranged by ascending median, that in the case of sentence length distributuons is a more suitable measure of central location than the mean, which is strongly affected by extreme values (WAKE 1957: 333; WILLIAMS 1969: 75).

Table 2 contains similar data for sentence length measured by number of syllables, while Table 3 covers sentence length by number of phonemes. Each Greek vowel, consonant and diphthong is counted as a single phoneme except the double consonants which are resolved into their parts i.e. ψ is counted as a π and a σ. The velar nasal, which has no special character in Greek, has been disregarded (ALLEN 1974: 33-37). In all cases, a sentence is a string of characters and blanks terminated by a strong stop, a weak stop, or an interrogation mark, which in modern texts are indicated by a low dot, high dot, and a semicolon.

A glance at Table 1 shows that in most cases, when sentence length is measured by number of words, the distribution peaks in the second category, sentences between 5 and 10 words in length, and descends slowly at first and then rapidly to zero. The *Argonautica*, the *Theogony*, *Theriaca*, and *Alexipharmaca*, which have longer sentences, have distributions that peak in the third category, sentences between 10 and 15 words in length. Sentence length distributions of Greek prose, measured by number of words and reported by WAKE and MORTON, have approximately the same shape, although the prose sentences are longer (WAKE 1957: 338ff; MORTON / LEVISON 1966: 156ff).

Figure 1 illustrates the distribution of sentence length calculated by number of words per sentence for all hexameter texts combined. Figure 1 may be compared with Figure 2 which shows the distribution of sentence length measured by number of syllables per sentence. In Figure 2, 4 separate peaks are visible. The highest peak is at the second category, sentences between 11 and 20 syllables. There is a dip at category 3 followed by a second rise in the fourth category (sentences between 31 and 40 syllables). In category 5 the distribution dips and then falls precipitously in 6. It rises modestly for a third time in category 7 (sentences between 61 and 70 syllables), dips again, and rises slightly for a fourth time in category 10 (sentences between 91 and 100 syllables). Table 2 indicates that several individual poems have this four-peaked distribution including the *Iliad, Odyssey, Theogony* and the *Hymn to Apollo*.Others have some of the peaks. The poem with a greater proportion of long sentences such as the *Theriaca* and the *Alexipharmaca* have peaks at categories 7 and 10 while poems with a greater proportion of shorter sentences have the peak at categories 2 and 4.

The peculiar multi-peaked sentence length distribution can be seen even more clearly when sentence length is calculated by number of phonemes. Figure 3 shows that there are clear peaks at categories 4 (sentences with 31 - 40 phonemes), 7 (61 - 70 phonemes), 10 (91 - 100 phonemes), 14 (131 - 140 phonemes), 17 (161 - 170 phonemes) and probably 20 although this category includes all sentences of 191 phonemes or more. Table 3 shows that when sentence length is measured by phonemes, the sentence length distribution of virtually every hexameter text has at least four good peaks, although in some cases the peak at category 10 is displaced to category 11, 13 to 14, and 16 to 17.

It is well to pause a moment to consider the reasons why a sentence length distribution of Greek hexameter poems has a series of four or more peaks at regular intervals. The answer lies in the metrical structure of the verse. The dactylic hexameter consists of six metrical feet any of which may have 2 or 3

syllables except the last which can have only 2. A dactylic hexameter line, then, has between 12 and 17 syllables. In fact, well over 70% of all hexameter lines have between 15 and 16 syllables (LA ROCHE 1898; 1899). Since the average syllable in Greek hexameter verse has 2.1 phonemes, most dactylic lines have between 31.5 and 33.6 phonemes. It follows then that all sentences exactly one line in length will be in category 2 (11 - 20 syllables) if sentence length is measured by syllables, and most will be in category 4 (31 - 40 phonemes) if sentence length is measured by phonemes. We may conclude, then, that the first and highest peak in the sentence length distribution is created by a disproportionate number of sentences which are exactly one line in length. A little arithmetic quickly shows that the second peak in the sentence length by phoneme distribution represents sentences exactly two lines in length, the third, three lines and so on. The peaks sometimes slip into contiguous categories because of the arbitrary way in which the categories are defined. For example, most four line sentences have between 60 and 64 syllables and fall neatly into category 7 of the sentence length by syllable distribution. These same sentences have between 126 and 134.4 phonemes, so that they are most often in category 14 of the sentence length by phoneme distribution, but occasionally in category 13.

It is clear, then, that the distribution of sentence length in Greek hexameter poetry reflects precisely the structure of the metrical line and illustrates how all Greek hexameter poets acknowledge that structure by making their sentences conform to its limits. When sentence length is measured by number of words, the relationship between verse end and sentence end is completely obscured. It is visible when sentence length is measured by syllables and is revealed most clearly when sentence length is measured by phonemes.

The poems with the sharpest peaks in the sentence length by phoneme distribution have the highest percentage of sentences ending at verse end. According to PARRY, the orally composed *Iliad* and *Odyssey* ought to have the highest proportion of sentencesending at verse end while the Hellenistic *Argonautica* ought to have a much lower proportion. To test this hypothesis, Table 4 was constructed, listing the name of each work, the number of its sentences which are punctuated at verse end and the percentage of its sentences punctuated at verse end. The works are listed by percentage from highest to lowest. Table 4 shows in effect how much attention an author gives to verse end in deciding where to end a sentence.

PARRY would be pleased to see that five of the six authors who pay least attention to verse end in punctuating their sentences are Hellenistic. The *Argonautica*, with only 57% of its sentences ending at verse end is last of all. In contrast, the *Iliad* and *Odyssey* have 74.5% and 75.9% sentences ending at verse end. So far PARRY's observations seem accurate. What PARRY could not learn from his partial study, however, is that an Hellenistic poem, Nicander's *Alexipharmaca*, has a proportion of sentence end coincident with verse end which at 78.5% is higher than those of the *Iliad* and *Odyssey*, and that two Homeric *Hymns*, the *Hymn to Apollo* and the *Hymn to Aphrodite*, have even higher proportions (80.7% and 84.4%). The two Homeric *Hymns* do not necessarily cause trouble for PARRY since there is evidence that they may be as old as the *Odyssey* (ALLEN et al. 1936: 183-186, 350-351), but the *Alexipharmaca* cannot be disposed of so easily. Whatever the merits of PARRY's theory we must concede what common sense surely approves, that a literate poet can end his sentences anywhere he wishes. He can even consciously imitate the sentence length patterns of oral poets.

What of sentences exactly one line in length? In this matter PARRY thought he saw a much sharper distinction between Homer and Apollonius. Table 5 lists each work together with the number of one line sentences it contains and the percentage from highest to lowest. Once again PARRY would be gratified to see that the *Iliad* and the *Odyssey* have the highest proportion of one line sentences (22% and 24% respectively) while four Hellenistic poems have the lowest proportions (*Phaenomena*, 10.9%; *Theriaca*, 6.0%; *Argonautica*, 5.8%; *Alexipharmaca*, 2.4%).

Once again, however, the list of poems is by no means chronological. Two other Hellenistic works, Callimachus' *Hymns* and the *Idylls*, have relatively high proportions of one line sentences while the *Theogony's* is relatively low.

These exceptions to PARRY's rule can be explained. Many of the *Idylls* are dialogues, a form that encourages the use of short sentences. Callimachus, with his wide knowledge of earlier poetry and superb technical skill, may have purposely imitated the style of the earlier *Homeric Hymns* which have proportions of one line sentences very close to that of his own *Hymns*. Furthermore, both Theocritus and Callimachus embraced an ideal of "slight" poetry which would seem to require short sentences rather than long ones. As for the *Theogony*, the subject matter and its disposition no doubt accounts for the relative lack of one line sentences in that poem. Unlike the *Iliad* and *Odyssey* which are narratives interspersed with speeches, the *Theogony* is "catalogue" poetry i.e. it contains a great many lists of names that cannot be accommodated by short one line sentences. In fact the longest sentence in all Greek hexameter poetry (508 phonemes = 118 words) is a list of the daughters of Tethys and Ocean, *Theogony* 346-361.

It seems, then, that PARRY is generally correct in his assumption that earlier poems, the products of a genuine oral tradition, contain a greater proportion of sentences ending at verse end and of sentences one line in length than later hexameter poems. There are important exceptions to this rule, however, among both early and later poems which show that subject matter, literary form, literary imitation and literary theory can all influence the location of sentence end even in so conservative a tradition as that of the Greek hexameter. While PARRY may be correct in linking the phenomena of the one line sentence and the coincidence of verse end and sentence end to the mechanics of oral composition, the relative quantities of these stylistic traits should not be considered reliable indications of the age of a poem or the manner in which it was composed.

We must now consider the matter of sentence length without regard to location of sentence end. Studies of sentence length in Greek prose lead us to expect that each author would have a sentence length distribution that is both homogeneous and unique i.e. that samples of at least 200 sentences taken from the sentences of a single author would belong to the same statistical population. Thus WAKE (1957) and MORTON / LEVISON (1966) use sentence length distributions to solve questions of disputed authorship.

Greek hexameter poetry provides interesting material with which to examine these notions. Scholars have often regarded large parts of the *Iliad, Odyssey, Theogony* and *Works and Days* as interpolations. Passages, indeed, whole books of the Homeric poems, have been rejected because they are allegedly inconsistent in style or spirit with the rest of the work or because they seem to disturb the poem's organization. West, for example, rejects everything after line 900 of the *Theogony*, but his is a conservative position. Mazon eliminates four passages totalling 225 lines, while Jacoby accepts as genuine only 442 lines of the 1022 in the standard text, (EDWARDS 1971: 4-5). The *Work and Days* has been subjected to similar treatment and the extensive depredations made by the "analysts" on the *Iliad* and *Odyssey* are well known, (DODDS 1968). Even the most conservative "unitarians" allow that the composition dates of some books such as book 10 of the *Iliad* and book 24 of the *Odyssey* are later than other parts of the text. In contrast, each of the Hellenistic hexameter poems is certainly the unique work of a single author.

Given that authors, or at least their individual works, have homogeneous, unique sentence length distributions, and given that some of the longer early hexameters have incorporated material of diverse origin, we would expect that the early hexameter poems would be less homogeneous in sentence length distribution than the Hellenistic poems. Statistics show that this assumption is wrong.

Let us consider first the Hellenistic poems. X^2 tests using samples of 200, 150 or 100 sentences, depending on the overall length of each work, show that some Hellenistic poems including Callimachus' *Hymns* (X^2 = 18.52, d.f. = 11; $X^2_{11;0.05}$ = = 19.7), the *Alexipharmaca* (X^2 = 2.47, d.f. = 7; $X^2_{7;0.05}$ = = 14.1) and the *Theriaca* (X^2 = 7.59, d.f. = 11; $X^2_{11;0.05}$ = 19.7) are homogeneous as we expect, but the *Idylls* (X^2 = 45.31, d.f. = = 14; $X^2_{14;\ 0.001}$ = 36.1), the *Argonautica* (X^2 = 64.44, d.f. = = 42; $X^2_{42;0.05}$ = 58.1) and the *Phaenomena* (X^2 = 47.97, d.f. = = 22; $X^2_{22;0.01}$ = 40.3) are not. In the case of the *Idylls*, the variegated nature of the poems themselves explains the inconsistency, but the *Argonautica* and *Phaenomena* are, on the surface, uniform in style and content.

Table 6 contains the relevant data on the *Argonautica*. When sentence length is measured by words, the sentence length distributions of the four books are homogeneous (X^2 = 31.97, d.f. = = 24; $X^2_{24;0.05}$ = 36.4), but when sentence length is measured by phonemes, they are not (X^2 = 64.36, d.f. = 42; $X^2_{42;0.05}$ = 58.1). Book 3 contains a smaller proportion of short sentences than books 1, 2, and 4, as well as some very long sentences. Books 1, 2 and 4 detail Jason's adventures at sea and in various landing places on his voyages to and from Colchis to recover the Golden Fleece. Book 3, in contrast, describes his stay in Colchis, dwelling particularly on his romantic relationship with the king's daughter, Medea. It seems then that the difference in sentence length distributions can be linked to a change in subject matter within the poem and that passion demands greater contrasts in sentence lengths than heroic adventure.

Changes in sentence length distributions can also be linked to changes in subject matter in the *Phaenomena*. The *Phaenomena* is unusually uniform and well organized as it is a versification of a prose work of the same name written by the astronomer Eudoxus of Cnidus. The subject matter and its treatment vary very little until line 733 where Aratus ceases describing the principal constellations and begins to discuss weather signs. This is hardly a major change, but it is worth noticing in view of the sentence length distributions. The sentences of the *Phaenomena* were divided into three samples and the data set out in Table 7. The first sample has the greatest proportion of shorter sentences, the second has a smaller proportion and the third, the least of all. The differences in the distributions are significant whether sentences length is measured by word (X^2 = 38.14, d.f. = 10; $X^2_{10;0.05}$ = 18.3) or phoneme (X^2 = 47.97, d.f. = 22; $X^2_{22;0.01}$ = = 40.3). As there is no likelihood of spurious additions, we can only assume that Aratus' style grew more full as he wrote. Since the third sample contains all of the sentences after 733, the change in subject at that point must have contributed to this developement.

While the sentence length distributions in Hellenistic poetry are less homogeneous than expected, the sentence length distributions of the early hexameter poems show surprising homogeneity. Both the *Theogony* (X^2 = 6.81, d.f. = 12; $X^2_{12;0.05}$ = = 21.0) and the *Works and Days* (X^2 = 11.36, d.f. = 10; $X^2_{10;0.05}$ = = 18.3) are uniform in sentence length distribution although both are thought to contain substantial interpolations. The *Shield of Heracles* (X^2 = 6.69, d.f. = 7; $X^2_{7;0.05}$ = 14.1), the *Hymn to Apollo* (X^2 = 4.22, d.f. = 7; $X^2_{7;0.05}$ = 14.1) and the *Hymn to Hermes* (X^2 = 13.48, d.f. = 8; $X^2_{8;0.05}$ = 15.5) are also homogeneous although the *Hymn to Demeter* (X^2 = 17.86, d.f. = = 9; $X^2_{9;0.05}$ = 16.9), the *Iliad* and the *Odyssey* are not. The *Hymn to Aphrodite* with only 135 sentences is too short for a test of homogeneity.

Considering the length of the *Iliad* and the *Odyssey* and the variety of subjects they treat, it is not surprising that the sentence length distributions of their many books are not uniform. The relevant data are in Tables 8 and 9. In the *Iliad* Books 1, 2, 7, 8, 12, and 23 are excluded from the population,

in the *Odyssey*, Books 5, 7, 11, 17, and 23. With the exception of *Odyssey* 11, none of these are the books most popularly excluded by the "analysts" who made their own judgments on other grounds. More interesting perhaps than the exclusion of individual books from a general population is the order in which the books fall when they are listed by ascending median as they are in Tables 8 and 9. We notice for example that the five books of the *Iliad* with the highest median sentence length, Books 15, 17, 20, 8 and 12, are all battle books with homogeneous sentence length distribution (X^2 = 53.63, d.f. = 48; $X^2_{48;0.05}$ = 65.2). *Il.* 13, another battle book with a somewhat lower median, is also within the limits of this population. In the *Odyssey*, Books 14-22 and 24 where Odysseus, disguised as a beggar, struggles to regain his throne after he has returned to Ithaca, represent 10 of the 14 books with the lowest median sentence length. Similarly, three of the four adventure books, Books 9, 11 and 12, the flashbacks in which Odysseus describes his fabulous voyage, are grouped together among those books with the highest median sentence lengths. Book 10, the fourth adventure book, has a much lower median. X^2 for all four adventure books, *Od.* 9-12, is 40.92 for 27 d.f. ($X^2_{27;0.05}$ = 40.1).

While groupings such as these suggest that subject matter and its disposition have an influence on sentence length distributions it is important to remember that it is only one of many possible factors. *Iliad* 21 e.g. is an aberant battle book with mean far below the others. One would have to seek beyond subject matter to see why this is so.

It is clear then that the suspect texts of early Greek hexameter poetry do not have less homogeneous sentence length distributions than the later Hellenistic poems. The tendency to end sentences at verse end and especially to make sentences one line long would seem to account for this consistency along with the standardization of other features of style. We could expect then that early hexameter poems dealing with similar subjects in a similar way would have identical or nearly identical sentence length distributions no matter who the author may be.

In one case at least, that of the four long *Homeric Hymns*, this hypothesis is clearly correct. In spite of differing authors and dates of composition the sentence length distributions of these four hymns all fall within the limits of a single population ($X^2 = 23.24$, d.f. = 27; $X^2_{27;0.05} = 40.1$). These poems have similar subject matter which they treat in a standard way. The Hellenistic *Hymns* of Callimachus share the same subjects and the same general format, but have a very different sentence length distribution.

A second group, the three older heroic epics, the *Iliad*, *Odyssey* and the *Shield of Heracles* have similar sentence length distributions, although there are statistically significant differences between the *Odyssey* and the others ($X^2 = 113.37$, d.f. = 24; $X^2_{24;0.001} = 51.2$). These three contrast plainly with the Hellenistic *Argonautica* just as the *Hymns* of Callimachus contrast with the *Homeric Hymns*. The direction of the contrasts is interesting. The *Hymns* of Callimachus have a much greater proportion of shorter sentences than the earlier *Hymns*. The *Argonautica* of Apollonius has a much greater proportion of long sentences than the earlier heroic poems. This difference may reflect a difference in literary ideals between Callimachus, who was a leading proponent of "slight" poetry, and Apollonius who is traditionally said to have opposed it,(PFEIFFER 1968: 142-144). In any case it shows that sentence length in Hellenistic hexameter poetry tends to greater extremes than in earlier poems of the same subgenre. The two hexameter works with the lowest median sentence length (Theocritus' *Idylls* and Callimachus' *Hymns*) and the two with the highest (Nicander's *Theriaca* and *Alexipharmaca*) are Hellenistic. As the earlier hexameter poems established the norms of subject matter, meter and diction which Hellenistic poets both imitated and developed, so too the early poems established norms in sentence length distribution which the Hellenistic poets acknowledged but also streched to suit their own purposes.

Summary

In conclusion we summarize our findings. When sentence length is measured by word, sentence length distributions in Greek hexameter poetry resemble those found in Greek prose by WAKE and MORTON, but when sentence length is measured by phoneme, multipeaked distributions are produced which illustrate how sentence end in hexameter poetry tends to coincide with verse end. Earlier poems, especially the *Iliad* and *Odyssey*, have a greater proportion of sentences ending at verse end and of sentences one line in length than later poems. These findings generally corroborate the incomplete statistics from which PARRY developed his theory linking the location of sentence end to the exigencies of oral composition. Exceptions to his rule, however, effectively forbid the use of sentence length distributions to date an hexameter poem or to determine whether or not it was composed orally.

Sentence length distributions are not always homogeneous within individual hexameter poems both early and late. When sentence length is measured by phoneme rather than by word, subtle variations in the sentence length distributions become apparent which can often be linked to changes in subject matter. Within a large poem like the *Iliad* or *Odyssey* books treating similar subjects in a similar way often have comparable sentence length distributions. Likewise whole poems of the earlier period which are similar in other major respects have similar sentence length distributions. Hellenistic hexameters which are given to stylistic extremes in all respects are a common exception to this rule. On the whole, when sentence length is measured by number of phonemes sentence length distribution is a sensitive index to varying style and subject matter, but an unreliable guide to date and authorship.

Table 1. Sentence Length Measured by Words in Greek Hexameter Poetry

	Number of Words Per Sentence																	
	1–5	6–10	11–15	16–20	21–25	26–30	31–35	36–40	41–45	46–50	51–55	56–60	61–65	66–70	Over 70	Total Sents.	Total Words	Medians
Idylls	321	472	325	148	91	35	23	13	12	3		1	1			1445	16751	9.8
Callimachus' *Hymns*	89	171	99	62	38	25	9	5	1	1	1	1	1			503	6345	10.3
Works and Days	69	172	104	67	39	17	5	4	2	1						480	5896	10.5
Shield of Heracles	32	88	58	34	20	5	5	6	2		1					251	3309	11.0
Odyssey	734	2472	1750	942	498	203	120	63	28	17	2	2	1			6832	87211	11.1
Iliad	763	2998	2221	1096	648	343	159	90	49	21	7	8			4	8407	111835	11.5
Hymn to Hermes	24	90	87	39	29	13	7	1	1		1					292	3987	12.3
Phaenomena	43	173	162	83	43	27	11	3	1	3	2					551	7602	12.3
Hymn to Demeter	29	64	47	30	29	10	8	2	2	2				1	1	225	3375	12.6
Hymn to Apollo	17	72	70	33	22	15	11	3	3		1				1	248	3811	13.0
Argonautica	246	675	694	482	247	146	64	26	33	8	4	2	1			2628	38853	13.3
Hymn to Aphrodite	10	36	35	24	16	5	5	1	1	1	1					135	2054	13.6
Theogony	36	105	121	71	43	24	5	11	6	4	1	1	1		3	432	6929	13.6
Theriaca	30	105	113	54	48	25	15	5	5	2						402	6209	14.3
Alexipharmaca	2	37	72	32	25	17	14	10	5	3		1				218	4096	16.4
All Greek Hexameters	2445	7730	5958	3197	1836	912	461	243	151	66	21	16	5			23049	308263	11.7

Table 2. Sentence Length Measured by Syllables in Greek Hexameter Poetry

	Number of Syllables Per Sentence																	
	1–10	11–20	21–30	31–40	41–50	51–60	61–70	71–80	81–90	91–100	101–110	111–120	121–130	131–140	Over 140	Total Sents.	Total Sylls.	Medians
Idylls	291	445	242	227	118	35	38	21	7	10	5	4		2		1445	36708	20.2
Works and Days	58	153	95	83	50	17	15	6	1	1	1					480	12940	23.6
Callimachus' *Hymns*	72	143	96	76	50	16	32	5	6	2	2		2		1	503	14613	24.3
Odyssey	635	2294	1187	1467	703	158	215	89	30	33	11	7	1	2		6832	190093	24.6
Iliad	619	2744	1592	1699	916	209	336	141	46	70	13	10	8	1	3	8407	246442	25.8
Shield of Heracles	28	67	54	48	25	10	3	9	1	4	1		1			251	7499	26.2
Phaenomena	29	127	141	107	73	26	23	12	7		3		3			551	17922	28.9
Hymn to Hermes	21	75	50	75	31	13	18	3	5			1				292	9289	30.5
Hymn to Apollo	16	67	34	55	37	7	14	11	1	3	2				1	248	8511	31.2
Hymn to Demeter	22	52	35	45	29	13	12	6	3	3	2	1			2	225	7846	31.3
Argonautica	156	516	575	509	393	171	146	62	31	39	14	8	6	1	1	2628	92212	31.8
Hymn to Aphrodite	8	32	23	34	16	6	7	6	1			1	1			135	4602	31.8
Theogony	20	91	79	98	69	17	25	5	6	13	3	1	1	1	3	432	16474	33.2
Theriaca	20	71	87	86	50	20	36	14	4	9	2	3				402	14935	33.2
Alexipharmaca	1	25	35	56	37	4	23	16	5	9		2	4		1	218	9949	39.1
All Greek Hexameters	1996	6902	4325	4665	2597	722	943	406	154	196	59	38	27	7	12	23049	690035	26.6

Table 3. Sentence Length Measured by Phonemes in Greek Hexameter Poetry

	Number of Phonemes Per Sentence																	
	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100	101-110	111-120	121-130	131-140	Over 140	Total Sents.	Total Phons.	Medians
Idylls	52	206	174	259	128	99	201	71	22	61	57	11	19	27	58	1445	78651	43.0
Callimachus' *Hymns*	14	53	56	85	52	48	57	14	17	29	17	7	14	17	23	503	30420	48.9
Works and Days	4	39	57	101	36	41	85	22	14	21	20	6	15	10	9	480	27554	51.2
Odyssey	46	486	662	1616	538	455	1336	366	167	423	236	66	115	119	201	6832	398920	51.2
Shield of Heracles	4	15	26	41	32	22	38	9	15	11	12	5	3	1	17	251	15940	53.9
Iliad	66	424	862	1834	701	686	1501	448	247	535	308	82	181	204	328	8407	518206	55.1
Phaenomena	2	23	38	81	62	62	98	26	31	46	17	13	12	14	26	551	37128	61.3
Hymn to Hermes	4	14	24	50	23	24	59	14	14	10	19	8	9	10	10	292	19341	61.7
Hymn to Demeter	7	15	11	39	19	16	29	12	12	14	15	4	6	8	18	225	16400	62.4
Hymn to Apollo	5	8	17	46	16	19	48	10	9	21	13	2	4	10	20	248	17913	63.2
Argonautica	67	73	206	251	290	290	336	189	152	183	158	81	88	74	190	2628	192490	64.6
Hymn to Aphrodite	1	6	12	21	7	7	32	10	3	11	6	1	6	2	10	135	9605	64.7
Theogony	5	13	33	54	28	30	93	28	17	34	26	6	15	15	35	432	33912	66.2
Theriaca	3	15	19	51	30	45	55	33	14	24	28	8	15	22	40	402	31898	67.4
Alexipharmaca	1		4	17	14	12	44	20	8	16	17	2	7	15	41	218	21355	79.0
All Greek Hexameters	281	1390	2201	4546	1976	1856	4012	1272	742	1439	949	302	509	548	1026	23049	1449733	56.6

Table 4. Sentences Punctuated at Line End in Greek Hexameter Poetry

	Sentences Punctuated at Line End	% Sentences Punctuated at Line End
Hymn to Aphrodite	114	84.4%
Hymn to Apollo	200	80.7%
Odyssey	5184	75.9%
Alexipharmaca	164	75.2%
Iliad	6268	74.6%
Theogony	314	72.7%
Hymn to Hermes	212	72.6%
Hymn to Demeter	163	72.4%
Works and Days	342	71.3%
Theriaca	267	66.4%
Idylls	933	64.6%
Callimachus' *Hymns*	325	64.6%
Phaenomena	351	63.7%
Shield of Heracles	147	58.6%
Argonautica	1498	57.0%

Note: The total number of sentences in each work is recorded in Tables 1, 2 and 3.

Table 5. Sentences One Line in Length in Greek Hexameter Poetry

	Number of One-Line Sentences	% of One-Line Sentences
Odyssey	1641	24.0%
Iliad	1850	22.0%
Works and Days	103	21.5%
Hymn to Apollo	53	21.4%
Idylls	273	18.9%
Hymn to Aphrodite	25	18.5%
Hymn to Hermes	52	17.8%
Callimachus' *Hymns*	88	17.5%
Hymn to Demeter	39	17.3%
Shield of Heracles	30	12.0%
Theogony	48	11.1%
Phaenomena	60	10.9%
Argonautica	152	5.8%
Theriaca	23	5.7%
Alexipharmaca	5	2.3%

Note: The total number of sentences in each work is recorded in Tables 1, 2, and 3.

Table 6. The Argonautica

(i)

Sentence Length Measured by Words

	Number of Words Per Sentence													
	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	Total Sents.	Total Words
Book 1	38	163	152	125	60	34	10	4	8	2	1	1	598	8976
Book 2	55	158	152	116	51	31	15	7	3	1	1		590	8508
Book 3	79	177	187	107	54	35	20	5	7	3	2		676	9601
Book 4	74	177	203	134	82	46	19	10	15	2	1	1	764	11768

(ii)

Sentence Length Measured by Phonemes

	Number of Phonemes Per Sentence																
	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100	101-110	111-120	121-130	131-140	Over 140	Total Sents.	Total Phons.
Book 1	6	14	41	56	64	59	93	39	42	50	37	17	24	17	39	598	44969
Book 2	15	16	53	60	60	60	71	50	31	50	37	15	16	21	35	590	42495
Book 3	25	23	60	76	78	84	87	37	31	42	35	22	12	13	51	676	46340
Book 4	21	20	52	59	88	87	85	63	48	41	49	27	36	23	65	764	58686

Table 7. The Phaenomena

(i)

Sentence Length Measured by Words

	Number of Words Per Sentence											
	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	Total Sents.	Total Words
Sample 1	24	80	55	23	11	4	2			1	200	2351
Sample 2	13	55	34	10	13	5	1	1		2	200	2868
Sample 3	7	38	39	28	20	8	6	1	1	2	150	2370

(ii)

Sentence Length Measured by Phonemes

	Number of Phonemes Per Sentence																
	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100	101-110	111-120	121-130	131-140	Over 140	Total Sents.	Total Phons.
Sample 1	1	12	18	34	31	24	32	9	10	12	4	3	1	4	5	200	10974
Sample 2	1	6	14	24	23	27	39	12	11	14	6	6	2	3	12	200	13441
Sample 3	0	5	6	23	8	11	27	5	9	20	7	4	9	7	9	150	11531

Table 8. Sentence Length Measured by Phonemes in the Iliad

Book	Number of Phonemes Per Sentence															Total Sents.	Total Phonemes	Medians
	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100	101-110	111-120	121-130	131-140	Over 140			
10	4	22	37	89	27	24	39	22	9	20	10	2	7	12	7	331	19194	45.5
1	4	16	46	90	21	20	70	15	10	17	11	3	5	8	12	348	20132	49.1
23	5	33	65	114	49	52	92	32	17	25	15	3	6	12	11	531	29743	50.4
22	6	15	36	62	19	19	50	16	8	18	10	1	6	1	15	282	17070	52.1
24	4	16	43	118	29	30	70	16	15	37	16	5	7	10	17	433	26593	52.7
18	5	23	41	65	32	40	56	20	10	14	18	6	5	4	11	350	20452	52.8
7	1	13	20	75	16	24	48	11	9	12	12	1	6	5	9	262	15728	53.0
4	2	11	35	77	21	22	55	17	9	18	16	5	6	5	7	306	17950	53.7
21	3	20	38	63	38	32	62	21	14	23	12	6	6	5	6	349	20313	54.4
3	2	20	22	62	13	12	49	7	6	17	10	2	12	8	6	248	15163	54.7
6	2	12	25	65	24	20	53	15	6	14	9	2	5	6	15	273	17390	54.8
9	6	26	33	73	34	31	51	23	17	28	15	4	6	11	15	373	23333	55.2
16	6	17	56	95	42	29	66	26	15	39	18	3	9	18	20	459	29014	55.2
19	2	6	24	50	22	20	45	13	6	12	8	1	3	10	8	230	14148	56.0
11	2	28	57	71	46	36	82	27	10	24	15	6	13	12	19	448	28057	56.1
14	1	12	26	67	21	21	52	18	4	17	13	4	6	7	9	278	17242	56.2
5		21	47	116	36	42	100	22	12	37	18	6	10	8	16	491	29966	56.6
13		27	47	78	47	35	73	26	19	33	15	4	9	7	21	441	27780	56.6
2	3	22	29	97	37	35	82	17	3	27	9	2	8	15	35	421	28596	56.9
15	3	16	33	76	32	36	67	21	13	23	12	7	11	13	16	379	24581	58.7
17	1	17	35	80	33	39	74	23	10	24	24	7	10	10	11	398	25140	59.0
20	2	10	24	50	25	24	53	19	9	20	7		6	6	10	265	16585	59.5
8	2	13	28	63	17	18	75	12	5	20	10	1	5	5	17	291	18451	61.1
12		8	15	38	20	25	37	9	11	16	5	1	14	6	15	220	15585	61.6

Table 9. Sentence Length Measured by Phonemes in the Odyssey

Book	Number of Phonemes Per Sentence: 1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100	101-110	111-120	121-130	131-140	Over 140	Total Sents.	Total Phonemes	Medians
17	5	30	44	103	26	24	69	21	9	16	11	6	3	4	6	377	19972	43.0
16	2	32	34	65	26	15	50	9	6	15	5	3	6	8	9	285	15751	44.2
19	5	25	29	108	18	26	58	19	8	25	17	3	5	4	6	356	19925	46.6
10	1	29	39	87	24	16	75	20	7	12	17		5	4	9	345	19099	47.4
20	1	25	22	54	20	17	48	10	7	8	4	3	3	4	8	234	13003	48.0
6	3	21	24	40	16	13	38	8	4	20	6		2	2	3	200	10790	48.0
2	2	15	25	61	24	15	50	7	4	17	6		6	5	8	245	14203	48.6
18	2	19	25	54	28	18	59	10	4	12	8	1	4	1	7	252	14083	49.8
14	3	27	27	69	28	24	58	15	8	18	8	7	5	5	7	309	17567	50.7
15	3	22	24	82	31	26	65	17	10	18	11	3	7	2	5	326	18298	51.0
21		19	20	55	25	17	45	10	7	11	8	2	8	6	7	240	14282	51.1
4	4	28	53	116	32	27	92	18	12	33	13	8	8	7	19	470	27914	51.2
24	4	23	33	68	25	22	56	21	9	15	11	3	4	5	11	310	18085	51.4
22		19	24	75	20	14	53	21	10	16	6	4	5	4	10	281	16731	52.3
8	1	24	28	79	28	20	66	19	3	22	12	1	9	6	10	328	19325	52.5
3	2	9	33	68	20	15	53	11	4	19	15	1	2	8	11	271	16443	52.8
1	1	16	22	61	13	12	40	13	9	14	7	2	7	6	11	234	14563	52.8
13	1	16	22	52	21	17	43	14	6	17	12	1	4	5	7	238	14393	54.6
12		14	19	58	19	23	51	13	4	16	11	7	2	5	5	247	14841	56.4
11		12	29	64	28	33	60	22	6	30	9	3	4	8	16	324	20947	59.3
9	1	22	27	61	28	20	74	20	11	22	8	2	8	6	5	315	18743	59.8
23	1	8	23	38	15	11	39	12	7	12	11	2	3	6	6	194	12231	60.8
5	2	20	19	58	16	18	56	24	9	21	11	3	1	4	8	270	16297	60.9
7	2	11	17	40	7	12	38	12	3	14	9	1	4	4	7	181	11434	60.9

FIGURE 1

GREEK HEXAMETER POETRY
SENTENCE LENGTH MEASURED BY NUMBER OF WORDS PER SENTENCE

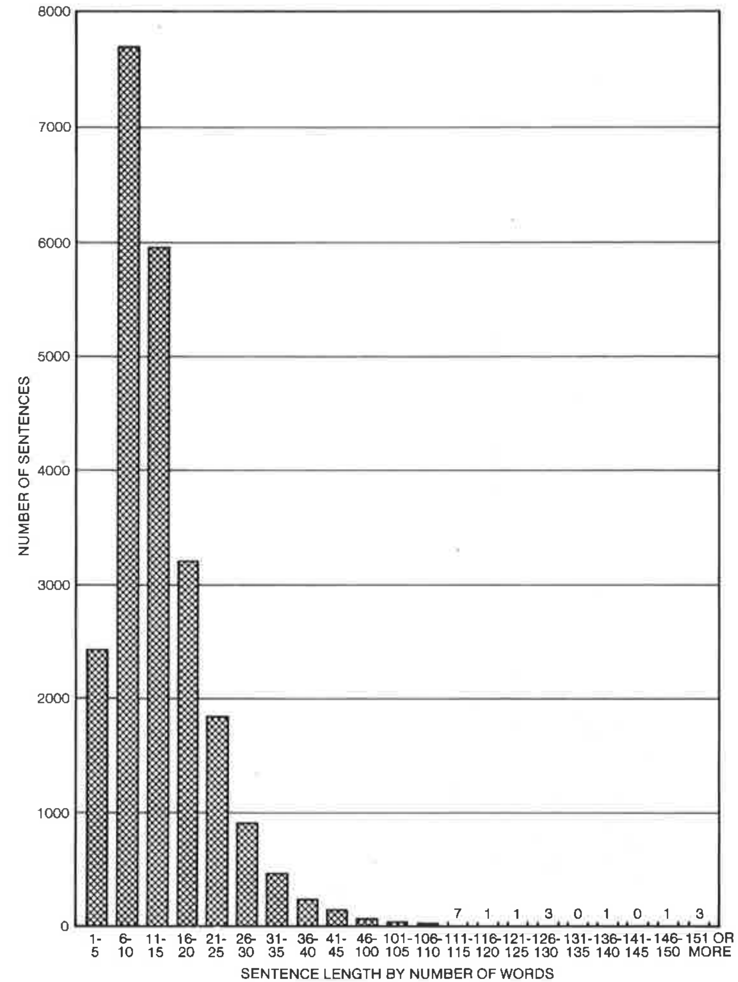


FIGURE 2

GREEK HEXAMETER POETRY
SENTENCE LENGTH MEASURED BY NUMBER OF SYLLABLES PER SENTENCE

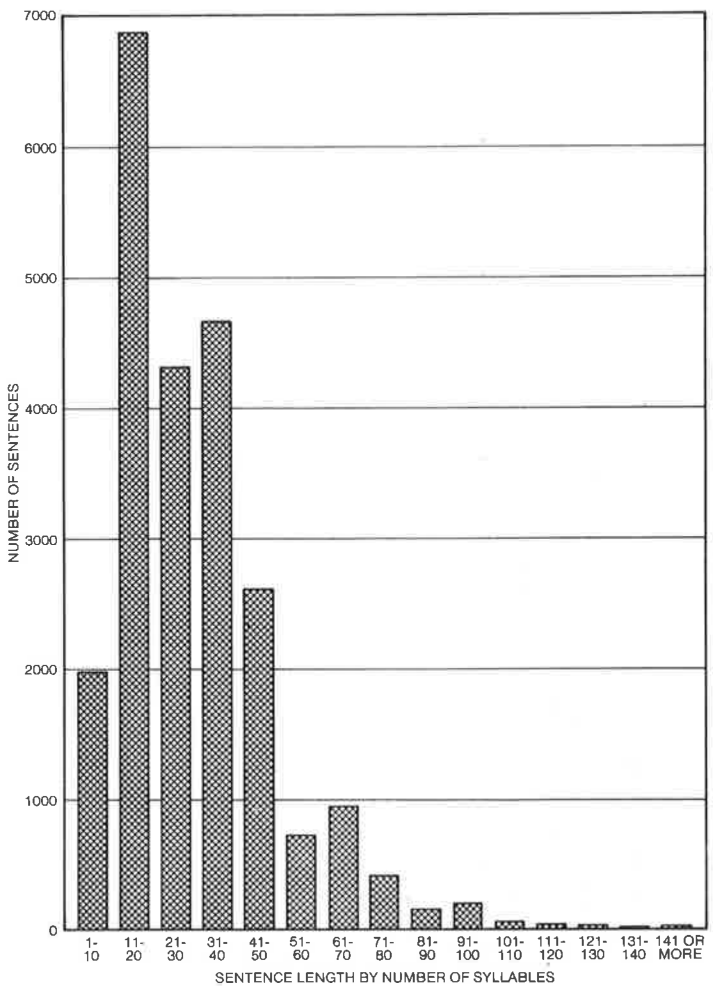
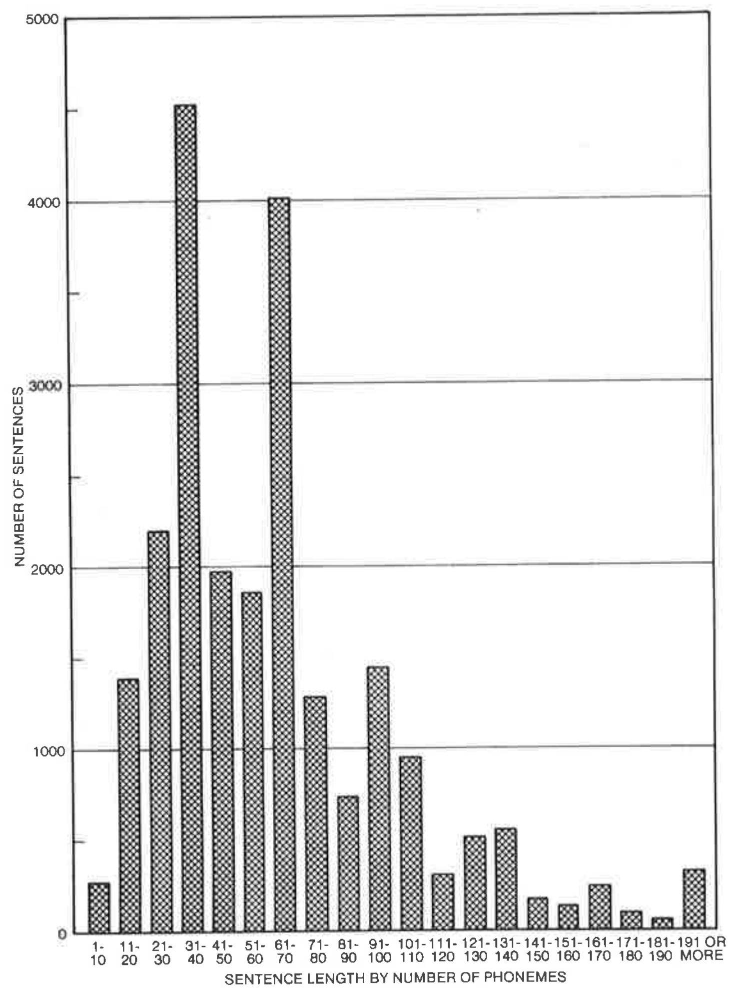


FIGURE 3

GREEK HEXAMETER POETRY
SENTENCE LENGTH MEASURED BY NUMBER OF PHONEMES PER SENTENCE



Notes

1 The author gratefully acknowledges the support of the American Philosophical Society, The American Council of Learned Societies, and the Board of Higher Education-Professional Staff Congress Research Award Program of the City University of New York. Nathan Mowatt, David Arnow, Armand Gazes and Malcolm Brown gave of their time and technical expertise. Those who generously contributed machine-readable texts are listed below in note 2. The author is indebted to all.

2 This study would not have been possible without the kind cooperation of the generous scholars who shared with the author machine readable texts prepared at great cost in time and labor:

Homer's *Odyssey*: P. Von der Mühll (Ed.), Editiones Helveticae. Text prepared by A.Q. Morton. Diacritics and capitals added by S.V.F. Waite with the aid of a text supplied by the Thesaurus Linguae Graecae.

Homer's *Iliad*: D.S. Monro and T.W. Allen (Eds.), *Homeri Opera*, Vols. 1 and 2 (Oxford 1920), reprinted 1962. Text prepared by A.Q. Morton. Diacritics and capitals added by D. Packard with the aid of a text supplied by the Thesaurus Linguae Graecae.

Hesiod's *Works and Days*, *Theogony*, *Shield of Heracles*: A. Rzach (Ed.), *Hesiodi Carmina* (Stuttgart 1967). Text prepared by C.A. Sowa. Diacritics and capitals added by J. Tebban.

The following texts were all prepared by the Thesaurus Linguae Graecae at the University of California, Irvine, T. Brunner, Director:

Callimachus' *Hymns*: R. Pfeiffer (Ed.), *Callimachus* Vol. 2 (Oxford 1949).

Aratus' *Phaenomena*: J. Martin (Ed.), *Arati Phaenomena* "La Nuova Italia" Editrice (Firenze 1956).

Apollonius' *Argonautica*: H. Fränkel (Ed.), *Apollonii Rhodii Argonautica* (Oxford 1961) reprinted 1967.

Theocritus' *Idylls*: A.S.F. Gow (Ed.), *Theocritus* (Cambridge 1950) reprinted 1965.

Nicander's *Alexipharmaca* and *Theriaca*: A.S.F. Gow and A.F. Scholfield (Eds.), *Nicander* (Cambridge 1950).

All texts were examined in their entirety with the following exceptions: In Callimachus' *Hymns* the elegiac fifth hymn was eliminated and two sentences containing significant lacunae were removed from *Hymn* 4, (4.171-185, 200-202). In Theocritus' *Idylls*: All elegiac passages were deleted as well as all titles, speakers' names, the poems considered by the editor to be spurious (*Idylls* 8, 9, 19, 20, 21, 23, 27), and the fragmentary poems.

References

ALLEN, W.S., Vox Graeca. Cambridge 1974

ALLEN, T.W./HALLIDAY, W.R./SIKES, E.E. (Eds.), The Homeric Hymns. Oxford 1936

CLAYMAN, D.L./VAN NORTWICK, T., Enjambement in Creek Hexameter Poetry. Transactions and Proceedings of the American Philological Association 107, 1978, 85-92

DODDS, E.R., Homer and the Analysts. In: PLATNAUER, M. (Ed.), Fifty Years (and Twelve) of Classical Scholarship. New York 1968, 1-8

EDWARDS, G.P., The Language of Hesiod in its Traditional Context. Oxford 1971

JANSON, T., The Problems of Measuring Sentence-Length in Classical Texts. Studia Linguistica 18, 1964, 26-36

LA ROCHE, J., Zahlenverhältnisse im Homerischen Vers. Wiener Studien 20, 1898, 1-69

LA ROCHE, J., Untersuchungen über den Vers bei Hesiod und in den homerischen Hymnen. Wiener Studien 20, 1898, 70-90

LA ROCHE, J., Der Hexameter bei Apollonios, Aratos und Kallimachos. Wiener Studien 21, 1899, 161-197

MORTON, A.Q./LEVISON, M., Some Indications of Authorship in Greek Prose. In: LEED, J. (Ed.), The Computer and Literary Style. Kent, Ohio 1966, 145-170

PARRY, M., The Distinctive Character of Enjambement in Homeric Verse. Transactions and Proceedings of the American Philological Association 60, 1929, 200-220, reprinted in PARRY, A. (Ed.), The Making of Homeric Verse. Oxford 1971, 251-265

PARRY, M., Whole Formulaic Verses in Greek and Southslavic Heroic Song. Transactions and Proceedings of the American Philological Association 64, 1933, 179-197

PFEIFFER, R., History of Classical Scholarship. Oxford 1968

WAKE, W.C., Sentence-Length Distribution of Greek Authors. Journal of the Royal Statistical Society, Series A 120, 1957, 331-346

WILLIAMS, C.B., A Note on the Statistical Analysis of Sentence-Length as a Criterion of Literary Style. Biometrika 31, 1939, 356-361

YULE, G.U., On Sentence-Length as a Statistical Characteristic of Style in Prose with Application to Two Cases of Disputed Authorship. Biometrika 30, 1938, 363-390

The Homogeneity of Metric Patterns in Hexameter

G. Altmann, Bochum

1. As far as we know the six metric feet of the hexameter line are not fully determined for any author but - at last as far as the first four feet are concerned - admit a free variation of two types of feet. In the majority of cases these are sequences of dactyls (D) and spondees (S). Since the metrical scheme is mainly determined by the first four feet we can disregard the last two; the following 16 patterns result

1. SSSS	5. DSSS	9. DSSD	13. DSDD
2. SSSD	6. SSDD	10. DSDS	14. DDSD
3. SSDS	7. SDSD	11. DDSS	15. DDDS
4. SDSS	8. SDDS	12. SDDD	16. DDDD.

It is neither the case that any author uses these patterns with equal frequency nor that any author tends to use only a single pattern. If all 16 patterns were used with equal frequency we would have *total homogeneity*; if only a single pattern were used we would have *total stereotypy*. Neither of these extremes has been observed. Nevertheless it is meaningful to ask *to what extent* the distribution of metric patterns in a poem is *homogeneous*.

Another kind of homogeneity is involved if we ask whether two authors use the individual patterns with equal frequency, i.e. whether they are similar in their use of metric patterns. One can compare for instance Homer with Horace, or Aeneis with Bucolica.

Finally, we can investigate the homogeneity of a whole group of authors or of the work of a single author. For instance one can examine the homogeneity of the complete works of Vergil or the works of all Latin authors.

Finally, we can examine whether two groups taken as wholes are homogeneous if compared with one another. One can for instance compare several Greek authors as a group with a group of German authors.

All these questions can be dealt with uniformly by means of *information statistics*. In the present paper we want to popularize the use of these simple methods rather than to make statements about individual authors or groups of authors. As far as we know the methods of information statistics have hardly ever been used in philology.

2. *The homogeneity of metric patterns in a single poem*

For the sake of simplicity let us designate the particular metric patterns with numbers 1, 2, ... , 16 as shown above and the frequencies of their occurrence in a poem with n_i (i = 1, 2, ... , 16) as shown in table 1. Here n is the number of examined lines, i.e. $n = \sum_{i=1}^{16} n_i$.

Table 1. Scheme of the distribution of metric patterns

Pattern	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Σ
Frequency	n_1	n_2	n_3	n_4	n_5	n_6	n_7	n_8	n_9	n_{10}	n_{11}	n_{12}	n_{13}	n_{14}	n_{15}	n_{16}	n

The usual method of computing the heterogeneity is the chi-square test according to the formula

$$X^2 = \sum_{i=1}^{k} \frac{[n_i - E(n_i)]^2}{E(n_i)} = \sum_{i=1}^{k} \frac{(n_i - \frac{n}{k})^2}{\frac{n}{k}} \qquad (1)$$

where k is the number of cells, here everywhere k = 16, and n is the sample size. Under the null hypothesis of homogeneity (equidistribution) one expects exactly n/k observations in each cell. X^2 is approximately distributed as χ^2 with k - 1 degrees of freedom.

By means of the likelihood-ratio test one can derive (cf. KENDALL, STUART 1967: 421) the test criterion

$$2I = 2\sum_{i=1}^{k} n_i \ln \frac{n_i}{E(n_i)} \qquad (2)$$

which by means of the Taylor series expansion of the logarithm (ignoring the members of order $O(n^{-1/2})$ leads to (1). The quantity (2) known as *minimum discrimination information statistics* (m.d.i.s.) is distributed approximately as χ^2 with k - 1 degrees of freedom and can be easily used for testing purposes (cf. KULLBACK 1959, KUPPERMAN 1960, KULLBACK, KUPPERMAN, KU 1962a, 1962b). Since we expect in each cell the same frequency $E(n_i)$ = n/k we obtain from (2)

$$2I = 2\sum_i n_i \ln \frac{n_i}{n/k} = 2\sum_i n_i \ln n_i - 2n \ln n + 2n \ln k. \qquad (3)$$

The computation consists merely of additions since extensive tables are available (cf. KULLBACK, KUPPERMAN, KU 1962b; BLÖSCHL 1966).

Example 1. DROBISCH (1868a) in a sample from the first book of Georgica by Vergil found the pattern frequencies which are presented in Table 2.

Table 2. A sample from Georgica by Vergil according to DROBISCH (1868a)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	n
31	18	16	55	84	11	19	29	39	62	49	9	19	31	29	12	513

We obtain

$$2\sum_i n_i \ln n_i = 2(31 \ln 31 + 18 \ln 18 + \ldots + 29 \ln 29 + 12 \ln 12) =$$
$$= 3749.0205$$
$$2n \ln n = 2(513 \ln 513) = 6402{,}5230$$
$$2n \ln k = 2(513 \ln 16) = 2844.6760,$$

from which

$$2I = 3749.0205 - 6402.5230 + 2844.6760 = 191.17.$$

Since $\chi^2_{.05(15)} = 25{,}0$ and the observed value of 2I is greater than 25,0 we reject the null hypothesis and consider the frequencies of metric patterns as not homogeneous. The author uses the patterns with different preference. For the sake of comparison: according to (1) we would abtain $X^2 = 207.44$.

3. *The homogeneity of two authors/texts*

Regardless of whether an author uses the patterns with the same frequency or not we can examine whether two authors use the same patterns to the same extent, i.e. whether the authors are "metrically homogeneous (similar)". The same holds for two texts of the same author.

Let us designate the frequency of the patterns as shown in Table 3.

Table 3. Scheme for the comparison of two authors

	1	2	3	... j	... 15	16	n_i
Author 1	n_{11}	n_{12}	n_{13} ...	n_{1j} ...	$n_{1,15}$	$n_{1,16}$	$n_{1\cdot}$
Author 2	n_{21}	n_{22}	n_{23} ...	n_{2j} ...	$n_{2,15}$	$n_{2,16}$	$n_{2\cdot}$
$n_{\cdot j}$	$n_{\cdot 1}$	$n_{\cdot 2}$	$n_{\cdot 3}$ ...	$n_{\cdot j}$ ...	$n_{\cdot,15}$	$n_{\cdot,16}$	n

In this case the test criterion is

$$2I = 2\sum_{i=1}^{2}\sum_{j=1}^{k} n_{ij} \ln \frac{n_{ij}}{E(n_{ij})}. \tag{4}$$

The expected frequency will be estimated here as

$$E(n_{ij}) = \frac{n_{i.}n_{.j}}{n}$$

from which

$$2I = 2\sum_{i=1}^{2}\sum_{j=1}^{k} n_{ij} \ln \frac{n_{ij}n}{n_{i.}n_{.j}} =$$

$$= 2\sum_{i=1}^{2}\sum_{j=1}^{k} n_{ij} \ln n_{ij} + 2n \ln n - 2\sum_{i=1}^{2} n_{i.} \ln n_{i.} -$$

$$- 2\sum_{j=1}^{k} n_{.j} \ln n_{.j}\,. \tag{5}$$

One should remember that $\sum_j n_{ij} = n_{i.}$ and $\sum_i n_{ij} = n_{.j}$ so that $\sum\sum_{ij} n_{ij} \ln n_{i.} = \sum_i n_{i.} \ln n_{i.}$ and $\sum\sum_{ij} n_{ij} \ln n_{.j} = \sum_j n_{.j} \ln n_{.j}$. The quantity 2I is distributed approximately as χ^2 with k - 1 degrees of freedom.

Example 2. Two samples from "Reinecke Fuchs" and "Hermann und Dorothea" by Goethe according to DROBISCH (1868b) yielded the values presented in Table 4. The computations according to (5) gave

Table 4. Samples from "Reinecke Fuchs" (the first three parts) and "Hermann und Dorothea" (the first five parts) by Goethe according to DROBISCH (1868b)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	$n_{i.}$
RF	14	7	96	204	6	33	86	181	5	46	93	71	22	42	65	43	1019
HD	15	11	40	200	31	35	42	149	25	34	104	98	25	73	63	41	1086
$n_{.j}$	29	18	136	404	37	73	228	330	30	80	197	169	47	115	128	84	2105

1014

$$2\sum_{ij}\sum n_{ij} \ln n_{ij} = 2(14 \ln 14 + 7 \ln 7 + \ldots + 63 \ln 63 + 41 \ln 41) = 19014{,}2095$$

$$2n \ln n = 2(2105 \ln 2105) = 32215{,}2178$$

$$2\sum_i n_{i.} \ln n_{i.} = 2(1019 \ln 1019 + 1086 \ln 1086) = 29299{,}2011$$

$$2\sum_j n_{.j} \ln n_{.j} = 2(29 \ln 29 + 18 \ln 18 + \ldots + 128 \ln 128 + 84 \ln 84) = 21841{,}9690$$

from which

$$2I = 19014{,}21 + 32215{,}22 - 29299{,}20 - 21841{,}96 = 88.26$$

The two samples are not to be considered homogeneous since $2I = 714{,}67 > 25{,}0 = \chi^2_{.05(15)}$.

Example 3. Comparing the patterns in the second part of Satires and in the first part of Epistulae by Horace we obtain according to DROBISCH (1868a) the frequencies presented in Table 5. Substituting the results in (5) we obtain

$$2I = 17925{,}04 + 31938{,}47 - 29045{,}34 - 20803{,}81 = 14{,}36.$$

Since $14{,}36 < 25{,}0$ we conclude that these two texts by Horace are metrically homogeneous.

Table 5. Satires (II) and Letters (I) by Horace according to DROBISCH (1868a)

S	83	43	72	109	153	36	50	52	67	103	106	32	41	55	64	17	1083
E	64	55	64	104	128	33	64	56	74	94	78	22	43	53	53	21	1006
$n_{.j}$	147	98	136	213	281	69	114	108	141	197	184	54	84	108	117	38	2089

4. *The homogeneity of several authors/texts*

The same test can be used for ascertaining the homogeneity of several (more than two) authors or texts. One can ask for instance whether the Iliad is homogeneous or whether there are "Vergilian" authors as assumed by DROBISCH (1866). Formula (5) can be directly used; the only difference consists in the number of degrees of freedom. If we have r authors then $i = 1,\ldots,r$ so that we have $(r-1)(k-1)$ degrees of freedom. The data should be arranged in the same way as shown in Table 6. The quantity

$$2I = 2\sum_{i=1}^{r}\sum_{j=1}^{k} n_{ij} \ln n_{ij} + 2n \ln n - 2\sum_{i=1}^{r} n_{i.} \ln n_{i.} - 2\sum_{j=1}^{k} n_{.j} \ln n_{.j} \tag{6}$$

is distributed approximately as χ^2 with $(r-1)(k-1)$ degrees of freedom.

Table 6. Scheme for comparing several authors/texts

	Metric patterns						
Author	1	2	...	j	...	k	Σ
1	n_{11}	n_{12}	...	n_{1j}	...	$n_{1,k}$	$n_{1.}$
2	n_{21}	n_{22}	...	n_{2j}	...	$n_{2,k}$	$n_{2.}$
...	...	...	...	...	...	...	...
i	n_{i1}	n_{i2}	...	n_{ij}	...	$n_{i,k}$	$n_{i.}$
...	...	...	...	...	...	...	...
r	n_{r1}	n_{r2}	...	n_{rj}	...	$n_{r,k}$	$n_{r.}$
Σ	$n_{.1}$	$n_{.2}$	...	$n_{.j}$	...	$n_{.,k}$	n

Example 4. It is to be tested whether Lucrece, Vergil, Horace, Manilius, Persius, Juvenal and Lucanus belong to the so

called Vergilian poets (the DROBISCH hypothesis). To this end we take from DROBISCH (1866) the corresponding samples as presented in Table 7. The method has the advantage that one can test several hypotheses at the same time using the same calculation: the homogeneity of metric patterns with individual authors, the homogeneity of individual authors with Vergil, or the homogeneity of the whole group. Here we restrict ourselves to the last problem. According to (6) we obtain

$$2\sum_{i=1}^{7}\sum_{j=1}^{16} n_{ij} \ln n_{ij} = 30277{,}16$$

$$2n \ln n = 66519{,}71$$

$$2\sum_{i=1}^{7} n_{i.} \ln n_{i.} = 50928{,}85$$

$$2\sum_{j=1}^{16} n_{.j} \ln n_{.j} = 45669{,}21$$

Table 7. Samples of Latin hexameter according to DROBISCH (1866)

1 Vergil	37	18	38	57	78	13	20	30	33	52	75	12	27	22	44	4	560
2 Horace	35	22	38	49	62	20	38	32	44	53	48	22	30	20	36	11	560
3 Lucrece	39	21	26	63	88	7	17	37	36	51	72	10	12	17	56	8	560
4 Manilius	48	11	34	67	93	16	22	28	33	60	57	9	19	30	22	11	560
5 Persius	30	12	27	62	118	5	35	19	48	68	96	9	32	35	39	14	649
6 Juvenal	25	13	38	67	85	14	26	22	40	64	51	20	19	29	35	12	560
7 Lucanus	15	12	32	58	83	8	23	28	39	98	59	11	19	29	33	13	560
$n_{.j}$	229	109	233	423	607	83	181	196	273	446	458	93	158	182	265	73	4009

from which

$$2I = 30277{,}16 + 66519{,}71 - 50928{,}85 - 45669{,}21 = 198{,}81.$$

The number of degrees of freedom is (7-1)(16-1) = 90. Since the observed value of 2I surpasses the critical value of $\chi^2_{.05(90)}$ =

113.1 we reject the null hypothesis of homogeneity and conclude that the concerned authors do not constitute a "Vergilian" group. Needless to say, there can exist a "Vergilian" group consisting of another combination of authors.

5. *The homogeneity of two groups*

In the same way it is possible to test whether two groups of authors are homogeneous. We start from the scheme as presented in Table 8. The group A of authors has the index i = 1, 2, ...,r, the group B the index i = r+1, r+2, ..., s. The particular marginal frequencies are defined as follows

$$n^A_{.j} = \sum_{i=1}^{r} n_{ij}$$

$$n^B_{.j} = \sum_{i=r+1}^{s} n_{ij}$$

$$n_{.j} = n^A_{.j} + n^B_{.j} = \sum_{i=1}^{s} n_{ij}$$

$$\left.\begin{aligned} n^A &= \sum_{j=1}^{k} n^A_{.j} \\ n^B &= \sum_{j=1}^{k} n^B_{.j} \end{aligned}\right\} \quad \text{here } k = 16$$

$$n = n^A + n^B = \sum_{\alpha=A}^{B} n^\alpha = \sum_{i=1}^{s} n_{i.} = \sum_{j=1}^{k} n_{.j}$$

Again, one can test the homogeneity within the groups, the homogeneity between the groups and the homogeneity of all authors. The formulas are given in Table 9. Writing ln(ab) = ln a + ln b we see that (7), (8) and (10) are identical with (6) and (9) is identical with (5). In formula (9) α is the variable of groups, i.e. α = A,B.

Table 8. Scheme for testing the homogeneity of two groups

		1	2	3	...	j	...	k	
A	1	n_{11}	n_{12}	n_{13}	...	n_{1j}	...	$n_{1,k}$	$n_{1.}$
	2	n_{21}	n_{22}	n_{23}	...	n_{2j}	...	$n_{2,k}$	$n_{2.}$
	...	...	...	...	...	...	...	...	...
	i	n_{i1}	n_{i2}	n_{i3}	...	n_{ij}	...	$n_{i,k}$	$n_{i.}$
	...	...	...	...	...	...	...	...	...
	r	n_{r1}	n_{r2}	n_{r3}	...	n_{rj}	...	$n_{r,k}$	$n_{r.}$
	$\sum$	$n^A_{.1}$	$n^A_{.2}$	$n^A_{.3}$	...	$n^A_{.j}$	...	$n^A_{.,k}$	n^A
B	r+1	$n_{r+1,1}$	$n_{r+1,2}$	$n_{r+1,3}$	...	$n_{r+1,j}$	...	$n_{r+1,k}$	$n_{r+1,.}$
	r+2	$n_{r+2,1}$	$n_{r+2,2}$	$n_{r+2,3}$	...	$n_{r+2,j}$	...	$n_{r+2,k}$	$n_{r+2,.}$
	...	...	...	...	...	...	...	...	...
	s	$n_{s,1}$	$n_{s,2}$	$n_{s,3}$	...	$n_{s,j}$	...	$n_{s,k}$	$n_{s.}$
	$\sum$	$n^B_{.1}$	$n^B_{.2}$	$n^B_{.3}$	...	$n^B_{.j}$	...	$n^B_{.,k}$	n^B
	$\sum$	$n_{.1}$	$n_{.2}$	$n_{.3}$	...	$n_{.j}$	...	$n_{.,k}$	n

Example 5. Let us examine whether there is homogeneity between Lucrece, Manilius and Lucanus on one hand and Voss and Goethe on the other. For the Latin authors we use the data from Table 7, for Voss the first two Idyls from "Luise" and for Goethe the first five parts of "Hermann und Dorothea" according to DROBISCH (1868b). The data are presented in Table 10. The values needed for the computation of m.d.i.s. are as follows

$$2\sum_{i=1}^{3}\sum_{j=1}^{16} n_{ij} \ln n_{ij} = 12708.47 \qquad 2\sum_{i=4}^{5}\sum_{j=1}^{16} n_{ij} \ln n_{ij} = 23080.09$$

$$2\sum_{i=1}^{3} n_{i.} \ln n_{i.} = 21261.87 \qquad 2\sum_{i=4}^{5} n_{i.} \ln n_{i.} = 36160.44$$

$$2\sum_{j=1}^{16} n^A_{.j} \ln n^A_{.j} = 16323.46 \qquad 2\sum_{j=1}^{16} n^B_{.j} \ln n^B_{.j} = 26136.32$$

$$2n^A \ln n^A = 24953.20 \qquad 2n^B \ln n^B = 39641.69$$

$$2\sum_{j=1}^{16} n_{.j} \ln n_{.j} = 47706.26$$

$$2n \ln n = 70229.36$$

From these numbers we obtain the results as given in Table 11. If zeros occur in the contingency table, i.e. if $n_{ij} = 0$ then we define

$$0 \ln 0 = 0$$

and *for each cell containing 0 we substract 1 from the resulting 2I* (cf. KU 1963). In our example only group B is concerned where in two cells with Voss zeros occur. The formula of 2I for homogeneity between A and B written explicitly yields

$$2I_{(AxB)} = 2\sum_{j=1}^{16} n^A_{.j} \ln n^A_{.j} + 2\sum_{j=1}^{16} n^B_{.j} \ln n^B_{.j} + 2n \ln n -$$
$$- 2n^A \ln n^A - 2n^B \ln n^B - 2\sum_{j=1}^{16} n_{.j} \ln n_{.j} = 414.99.$$

All computed values in Table 11 are very high. The critical values of χ^2 at $\alpha = 0.05$ are

$$\chi^2_{(15)} = 25.0$$

$$\chi^2_{(30)} = 43.8$$

$$\chi^2_{(60)} = 79.1$$

Table 9. Formulas for testing the homogeneity

Component	Information	Degrees of freedom	
Homogeneity within A	$2 \sum_{i=1}^{r} \sum_{j=1}^{k} n_{ij} \ln \frac{n^A n_{ij}}{n_{i.} n^A_{.j}}$	$(r-1)(k-1)$	(7)
Homogeneity within B	$2\sum_{i=r+1}^{s} \sum_{j=1}^{k} n_{ij} \ln \frac{n^B n_{ij}}{n_{i.} n^B_{.j}}$	$(s-r-1)(k-1)$	(8)
Homogeneity between A and B	$2\sum_{\alpha=A}^{B} \sum_{j=1}^{k} n^{\alpha}_{.j} \ln \frac{n n^{\alpha}_{.j}}{n^{\alpha} n_{.j}}$	$(k-1)$	(9)
Homogeneity total	$2 \sum_{i=1}^{s} \sum_{j=1}^{k} n_{ij} \ln \frac{n n_{ij}}{n_{i.} n_{.j}}$	$(s-1)(k-1)$	(10)

Table 10. Samples from Latin and German authors

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1 Lucrece	39	21	26	63	88	7	17	37	36	51	72	10	12	17	56	8	560
2 Manilius	48	11	34	67	93	16	22	28	33	60	57	9	19	30	22	11	560
3 Lucanus	15	12	32	58	83	8	23	28	39	98	59	11	19	29	33	13	560
$n^A_{.j}$	102	44	92	188	264	31	62	93	108	209	188	30	50	76	111	32	1680
4 Voss	0	0	32	51	9	35	154	131	44	95	70	188	173	164	135	161	1442
5 Goethe	15	11	40	200	31	35	142	149	25	34	104	98	25	73	63	41	1086
$n^B_{.j}$	15	11	72	251	40	70	296	280	69	129	174	286	198	237	198	202	2528
$n_{.j}$	117	55	164	439	304	101	358	373	177	338	362	316	248	313	309	234	4208

161

Table 11. The computation of homogeneity for the data in Table 10

Component	Information	Degrees of freedom
Homogeneity within A	76.34	30
Homogeneity within B	396.02	15
Homogeneity between A and B	414.99	15
Homogeneity total	887.35	60

We can conclude that neither A nor B nor both with one another nor all the five authors with one another are to be considered as homogeneous.

The method can be extended to any number of groups and can be used for the solution of several problems: characterisation, comparison, classification, measurement of evolutionary tendencies, "imitation" of one author by an other author etc. Other kinds of partitions can be found in KULLBACK (1958).

References

BLÖSCHL, L., Kullbacks 2I-Test als ökonomische Alternative zur chi^2-Probe. Psychologische Beiträge 9, 1966, 379-406.

DROBISCH, M.V., Ein statistischer Versuch über die Formen des lateinischen Hexameters. Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig. Philologisch-historische Classe 18, 1866,73-139.

DROBISCH, M.V., Weitere Untersuchungen über die Formen des Hexameters des Vergil, Horaz und Homer. Berichte ü.d.Verh.d.Kön. Sächs.Ges.d.Wiss. zu Leipzig. Phil.-hist. Classe 20, 1868a, 16-53.

DROBISCH, M.V., Über die Formen des deutschen Hexameters bei

Klopstock, Voss und Goethe. Berichte ü.d.Verh.d.Kön.Sächs. Ges.d.Wiss. zu Leipzig. Phil.-hist. Classe 20, 1868b, 138-160.

KENDALL, M.G., STUART, A., The advanced theory of statistics. Vol. 2. London, Griffin 1967.

KU, H.H., A note on contingency tables involving zero frequencies and the 2I test. Technometrics 5, 1963, 398-400.

KULLBACK, S., Information theory and statistics. New York, Wiley 1959.

KULLBACK, S., KUPPERMAN, M., KU, H.H., Tests for contingency tables and Markov chains. Technometrics 4, 1962a, 573-608.

KULLBACK, S., KUPPERMAN, M., KU, H.H., An application of information theory to the analysis of contingency tables, with a table of 2n ln n, n = 1(1)10,000. Journal of Research, National Bureau of Standards, Section B, 66B, 1962b, 217-243.

KUPPERMAN, M., On comparing two observed frequency counts. Applied Statistics 9, 1960, 37-42.

Vocalic Initials in Aeneid 12

Nathan A. Greenberg, Oberlin

In the following study, we consider closely the distribution of vocalic initials in Latin hexameter verse. In the course of our consideration, we isolate some anomalous aspects of that distribution and attempt some explanation for these. We use *Aeneid* 12 as a paradigm case, but much of what is seen here is true of the Latin hexameter verse in general.

The text of *Aeneid* 12, as analyzed by OTT (1973), consists of 949 complete hexameter verses. OTT[1] (p. 66) also lists the number of junctures *(Wortgrenzen)* and elisions occurring at the sixteen possible positions within the verse. To his listing we add in Table I the number and percentage of words beginning at these positions with a vowel or with the letter *'h'*. (For completeness, we also add a Position 0 to indicate the beginning of the verse.)

Table I supplies some interesting information. It has, of course, been long observed that the frequency of junctures varies widely within the Latin hexameter, in this case from 82.02% at Position 7 (after the first syllable in Foot 3) to 0.74% at Position 16 (after the first syllable in Foot 6). The observation has been accompanied by speculations accounting for this unequal distribution. STURTEVANT (1921), for example, stated, we believe correctly, that the obvious avoidance of juncture at Positions 13 and 16 (after the first syllables in Feet 5 and 6) was due to the desire, need, or necessity of maintai-

Table I. *Aeneid* 12: Frequency of All Initials, of Elisions, and of Vocalic Initials at Each Position in the Verse.

Position	(a) All Initials	(b) (a)/949	(c) Elisions	(d) (c)/(a)	(e) Vocalic Initials	(f) (e)/(a)
0	949	1.0000			363	0.3825
1	321	0.3383	61	0.1900	99	0.3084
2	155	0.1633	5	0.0323	52	0.3355
3	397	0.4183	98	0.2469	269	0.6776
4	631	0.6649	31	0.0491	98	0.1553
5	129	0.1359	2	0.0155	25	0.1938
6	156	0.1644	66	0.4231	99	0.6346
7	780	0.8219	62	0.0795	177	0.2269
8	149	0.1570	1	0.0067	23	0.1544
9	177	0.1865	51	0.2881	80	0.4520
10	724	0.7629	103	0.1423	225	0.3108
11	36	0.0379	1	0.0278	14	0.3889
12	522	0.5501	37	0.0709	150	0.2874
13	19	0.0200	3	0.1579	4	0.2105
14	449	0.4731	6	0.0134	114	0.2539
15	550	0.5796	1	0.0018	218	0.3964
16	7	0.0074	0	0.0000	0	0.0000
Totals	6151		528	0.0858	2010	0.3268

ning the consonance of metrical ictus and word accent in the final two feet of the verse[2].

With regard to elisions, we may be struck by their paucity (both in absolute number and in percentage) at Positions 2, 5, 8, 11, 14 (i.e., between the two short syllables of the dactyl)

(cf. SOUBIRAN 1966: 207 ff.).

Our central concern, however, is with vocalic initials. The findings in Column (e) in Table I are our addition to OTT. (They can be derived from the data in OTT, but only with considerable difficulty.) Of the 6151 words[3] in *Aeneid* 12, 2010 (or 32.68%) begin with a vowel or with the letter '*h*'. There are, of course, constraints upon the deployment of initials in the hexameter line. It is, therefore, interesting to note that at the beginnings of verses, where there are no constraints of position, the percentage of vocalic initials (henceforth to be understood as including initial '*h*') is 38.25%. SOUBIRAN has a good discussion of disparities of this sort. His conclusion, with which we agree, is that, since the normal rate of incidence of initial vowels is about 35%, the lower rate of incidence within the verse is due to a general avoidance of elision, which includes both a general avoidance of words with vocalic initials and some displacement of such words to the beginning of the verse. Some comparative data are available in Table II. With the exception of the *Ars Poetica* and the third book of Lucretius, a notably larger percentage of the first words of the verse have vocalic initials than do the words as a whole (cf. SOUBIRAN 1966: 590 ff)[4].

Our own interest is focussed upon the fact, as shown in Column (f) of Table I, that there are three positions within the verse, Positions 3, 6, 9 (following the final syllables of Feet 1, 2, 3 respectively), where the percentage of vocalic initials is significantly higher than that found at the beginning of the verse[5]. In particular, over 60% of the junctures at Positions 3 and 6 have vocalic initials. Further, Position 3 has the lar-

Table II. Comparative Frequencies of All Vocalic Initials and of Vocalic Verse Initials in Selected Works (Raw Data in Parentheses).

	Percentage of all words with vocalic initials		Percentage of vocalic verse initials	
Catullus 64	23.15	(558/2410)	29.41	(120/ 408)
Lucretius, Bk. 3	33.30	(2443/7336)	32.91	(359/1091)
Culex	29.53	(768/2601)	41.84	(172/ 411)
Ecloques	29.52	(1660/5624)	36.87	(306/ 830)
Aeneid 4	31.43	(1435/4565)	35.60	(251/ 705)
Aeneid 12	32.68	(2010/6153)	38.25	(363/ 949)
Ars Poetica	30.06	(926/3081)	29.41	(140/ 476)
Metamorphoses 12	31.03	(1265/4077)	36.94	(232/ 628)
Lucan, Bk. 1	24.26	(1055/4348)	36.23	(250/ 690)
Lucan, Bk. 10	25.77	(906/3516)	37.77	(207/ 548)
Valerius Flaccus, 1	28.36	(1590/5606)	36.51	(310/ 849)

gest number (let alone percentage) of vocalic initials. The question we wish to explore is: what is so peculiar about Position 3 that vocalic initials tend to cluster there to such a degree?

It is also noteworthy, as we see from Column (d) in Table I that Positions 3, 6, 9 are also the places where the largest percentages of elisions are found. It seems reasonable to conclude that there may be some relation between high frequencies of vocalic initials and elisions. After all, one cannot have elision without a vocalic initial[6]. One can, however, have vocalic initials without elisions, and these are the cases which are explored below.

Further, since Positions 3, 6, 9, 12 all occur at the ends of metrical feet, variations in quantity come into consideration. When Foot 1 is a dactyl rather than a spondee, Position 3 occurs at the end of a short syllable. In such a case, either the previous word ends with a single consonant and the following word has a vocalic initial or the previous word ends with a short vowel and the following word begins with a consonant. (The remaining possibilities are elision and hiatus which we are eliminating from this consideration.) It is obvious, then, that vocalic initials are necessarily present in one of the two ways in which juncture may occur at the end of a dactyl. By the same token, vocalic initials cannot be present when the other mode of juncture occurs. Since the same constraints do not apply to junctures occurring at the ends of spondaic feet, it has seemed desirable to us to separate the data for dactylic and spondaic feet. This is done in Table III.

In Table III, data are listed for Positions *d3* (occurring at the end of a Foot 1 dactyl) and *s3* (occurring at the end of a Foot 1 spondee), and so on for Positions 6, 9, and 12. (It has not seemed necessary to subdivide the observations for Position 15, at the end of Foot 5, since the book contains only two spondaic verses.) Thus we see in Column (a) of Table III that of the 397 junctures at Position 3, 269 occur after a dactyl in Foot 1 and 128 after a spondee in that foot. The disparity is not unexpected since there are in general more dactylic than spondaic first feet.

It is much more startling to note, by a comparison of Columns (c) and (e) in Table III, that of the 63 vocalic initials occurring after a spondee in Foot 1 no less than 56 are involved in elision. Just so, 41 out of 48 at Position *s6* (after a

Table III. *Aeneid* 12: Frequencies and Percentages of Junctures, Elisions, and Vocalic Initials with Differentiation at Positions 3, 6, 9, 12.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Position	All. Inits.	(a)/949	Elisions	(c)/(a)	Vocalic Inits.	(e)/(a)	(e)-(c)	(g)/ (a)-(c)
0	949	1.0000			363	0.3825	363	0.3825
1	321	0.3383	61	0.1900	99	0.3084	38	0.1462
2	155	0.1633	5	0.0323	52	0.3355	47	0.3133
d3	269	0.2835	42	0.1561	206	0.7658	164	0.7225
s3	128	0.1349	56	0.4375	63	0.4922	7	0.0972
4	631	0.6649	31	0.0491	98	0.1553	67	0.1117
5	129	0.1359	2	0.0155	25	0.1938	23	0.1811
d6	62	0.0653	25	0.4032	51	0.8226	26	0.7027
s6	94	0.0991	41	0.4362	48	0.5106	7	0.1321
7	780	0.8219	62	0.0795	177	0.2269	115	0.1602
8	149	0.1570	1	0.0067	23	0.1544	22	0.1486
d9	50	0.0527	21	0.4200	43	0.8600	22	0.7586
s9	127	0.1338	30	0.2362	37	0.2913	7	0.0722
10	724	0.7629	103	0.1423	225	0.3108	122	0.1965
11	36	0.0379	1	0.0278	14	0.3889	13	0.3714
d12	134	0.1412	6	0.0448	84	0.6269	78	0.6094
s12	388	0.4089	31	0.0799	66	0.1701	35	0.0980
13	19	0.0200	3	0.1579	4	0.2105	1	0.0625
14	449	0.4731	6	0.0134	114	0.2539	108	0.2438
15	550	0.5796	1	0.0018	218	0.3964	217	0.3953
16	7	0.0074	0	0.0000	0	0.0000	0	0.0000
Totals	6153		528	0.0858	2010	0.3268	1482	0.2635

spondaic Foot 2) are involved in elision, and so also 30 out of 37 vocalic initials at Position *s9* (after a spondaic Foot 3)[7]. When the effects of elisions are removed, as in Columns (g) and (h) of Table III, it is seen that very small numbers and percentages of vocalic initials remain at Positions *s3*, *s6*, *s9*, and *s12*. By contrast, it is clearer than ever from Column (h) in Table III that significantly large percentages of vocalic initials occur at Positions *d3*, *d6*, *d9*, and *d12* even when junctures with elision are eliminated. The percentages range from 75.86% at *d9* (after a dactylic Foot 3) to 60.94% at *d12* (after a dactylic Foot 4).

We might shrug our shoulders and simply accept these high percentages as characteristic of those words in the language which satisfy the constraints present at junctures following a dactylic foot, were it not for the strongly contrasting data at Position 15 (at the end of Foot 5, the foot which is more dactylic than any other). Here we find a surprising 39.53%, much lower than the percentages we have seen at the ends of the other dactylic feet.

In order to examine more closely how this disparity comes about, we have chosen to make a census of the metrical word-types ending at Positions *d3*, *d6*, *d12*, and 15. (It is, after all, the metrical shape of the preceding word that is affected by a vocalic initial.) In Tables IV and V, we list the number of each word-type and the number and percentage of these ending in a consonant. The number and percentage of following vocalic initials will, of course, be the same. Table IV presents the census by positions, Table V by word-types.

From Table IV, it can be seen that we are dealing with a very limited number of word-types at Positions *d3*, *d6*, and *d9*.

Table IV. *Aeneid* 12: Metrical word-types ending unelided at Positions *d3*, *d6*, *d9*, *d12*, and 15 by position.

Ending at Position	Word-type	(a) Number	(b) Number with final consonant	(c) (b)/(a)
d3	u	22	19	0.8636
	uu	78	52	0.6667
	-uu	127	93	0.7323
	Total	227	164	0.7225
d6	u	1	1	1.0000
	uu	36	25	0.6944
	Total	37	26	0.7027
d9	u	3	3	1.0000
	uu	26	19	0.7308
	Total	29	22	0.7586
d12	u	12	12	1.0000
	uu	84	52	0.6190
	-uu	18	8	0.4444
	u-uu	3	2	0.6667
	uu-uu	3	2	0.6667
	--uu	8	2	0.2500
	Total	128	78	0.6093
15	u	57	57	1.0000
	uu	10	1	0.1000
	-uu	337	108	0.3205
	u-uu	13	5	0.3846
	uu-uu	7	3	0.4286
	--uu	120	41	0.3417
	---uu	3	1	0.3333
	-uu-uu	1	0	0.0000
	Total	548	216	0.3942

Table V. *Aeneid* 12: Metrical word-types ending unelided at Positions *d3*, *d6*, *d9*, *d12*, and 15 by word-type.

Word-type	Ending at Position	(a) Number	(b) Number with final consonant	(c) (b)/(a)
u	*d3*	22	19	0.8636
	d6	1	1	1.0000
	d9	3	3	1.0000
	d12	12	12	1.0000
	15	57	57	1.0000
	Total	95	92	0.9684
uu	*d3*	78	52	0.6667
	d6	36	25	0.6944
	d9	26	19	0.7308
	d12	84	52	0.6190
	15	10	1	0.1000
	Total	234	149	0.6367
-uu	*d3*	127	93	0.7323
	d12	18	8	0.4444
	15	337	108	0.3205
	Total	482	209	0.4336
u-uu	*d12*	3	2	0.6667
	15	13	5	0.3846
	Total	16	7	0.4375
uu-uu	*d12*	3	2	0.6667
	15	7	3	0.4286
	Total	10	5	0.5000
--uu	*d12*	8	2	0.2500
	15	120	41	0.3417
	Total	128	43	0.3359

Indeed, Positions *d6* and *d9* (not very numerous in any case) are preceded only by monosyllables and pyrrhics (uu), and we deal with these word-types first.

Monosyllables play no role in the differentiation we seek. Aside from the single occurrence before Position *d6*, monosyllables constitute about one tenth of the total occurrences at each position, and almost all of them end with a consonant[8].

The case with regard to pyrrhics (uu) is more interesting. They constitute the overwhelming majority of the words ending at Positions *d6* and *d9*. They also constitute about 65% of the words ending at Position *d12*. In general, as we see from Table V, about 64% of these words end in a consonant, and this is true of their occurrences in all the first four feet. They are largely responsible for the preponderance of vocalic initials at Positions *d6*, *d9*, and even *d12*, and they make an important contribution to Position *d3*. More noteworthy is their virtual absence from Foot 5, where there are only ten out of a total of 548 junctures. As it happens, only one of the ten ends in a consonant, but it makes little difference statistically whether all or none of them do so. It is more important to note that by their absence from Foot 5 (assuming that as elsewhere about 65% of such words would end in a consonant), they allow the foot to have a larger proportion of final vowels than it might otherwise have. At the same time, we do not believe that this is why such words are avoided in Foot 5. Rather, as we see from OTT (1973: 120), such words occur in Foot 5 only when preceded by a monosyllable. For example,

semianimi lapsoque supervenit, et pede collo

(*Aeneid* 12.356).

Monosyllables at this position in the verse are very rare[9]. One

may surmise that a monosyllable is required in such a case in order to preserve the consonance of ictus and accent in Foot 5, but the use of a following pyrrhic must of necessity set up a strongly competing accent in the second syllable of the foot. It is hard to believe that the accent on *et* in the above example can compete with the stress laid upon the first syllable of *pede*.

Most interesting are the differences seen in the distribution of dactylic words. As we see from Table V, these occur in significant numbers only in Feet 1 and 5, where they constitute well over half the occurrences of words ending at those positions. About 73% of the dactylic words in Foot 1 end in a consonant, thereby requiring a vocalic initial in the following word. In very strong contrast, only 32% of the dactylic words in Foot 5 end in a consonant, and so only that percentage of the following words have vocalic initials. We believe that the key to our question lies in this last clause. Well over half the verses of *Aeneid* 12 (about 58%) end in bisyllabic words. These final words constitute one of the very noticeable recurring features of the verse. It would be, we believe, at the very least inconvenient and perhaps aesthetically jarring, if this large class of words were to have a noticeably larger proportion of voaclic initials than the language as a whole. We have already noted above that about 33% of the words in *Aeneid* 12 have vocalic initials. We now add the observation that about 52% of all the words in *Aeneid* 12 end with a vowel or with the letter '*m*'[10]. As we see from Table V, of the 482 dactylic words in *Aeneid* 12 about 57% end with a vowel, not a very great disparity. In order to bring about a pleasing proportion of initials for the large set of final bisyllabic words, Vergil places in Foot 5 a rather larger proportion of dactylic words with vocalic finals than would be expected. In turn, a larger proportion than would be expected of dactylic words with consonantal finals is displaced to the only place where they can go in sizable numbers, i.e., to the first foot. These must be followed by vocalic initials and this largely accounts for the unusually large number and percentage of vocalic initials at Position *d3*. Our finding is not confined to Vergil. As we see from the comparative data displayed in Table VI, the displacement seen in *Aeneid* 12 is encountered to a greater or lesser degree in all the works surveyed.

Table VI. Comparative Percentages of Vocalic Initials after Unelided Dactylic Feet at Positions *d3* and 15 in Select Works. (Raw Data in Parentheses)

	Position *d3*		Position 15	
Catullus 64	46.15	(60/130)	18.37	(45/245)
Lucretius, Bk. 3	58.17	(203/349)	48.78	(260/533)
Culex	53.28	(73/137)	44.81	(121/270)
Ecloques	52.33	(146/279)	37.20	(186/500)
Aeneid 4	66.67	(120/180)	38.03	(143/376)
Aeneid 12	72.25	(164/227)	39.42	(216/548)
Ars Poetica	63.57	(82/129)	46.47	(125/269)
Metamorphoses 12	61.99	(181/292)	45.39	(182/401)
Lucan, Bk. 1	70.04	(159/227)	39.95	(159/398)
Lucan, Bk. 10	66.05	(107/162)	43.42	(132/304)
Valerius Flaccus, 1	71.34	(219/307)	35.59	(200/562)

There remains one more word-type, the Ionic a majore (--uu), which terminates in significant numbers at Position 15. The word-type is rare elsewhere, so there can be no conjectures about displacements. One can only say that Vergil's choice of

these words has been such as to result in only 34% of consonantal finals, very close to the desired proportion.

These last anticlimactic remarks should serve as a humbling caveat for the assumptions of the preceding paragraphs. It has not been our intention to suggest that Vergil was in any way trammeled or hindered by the formal requirements of the hexameter verse. Rather, it seems better to conclude that in some sense he responded positively to these formal demands, that they constituted a creative challenge that the consummate poet took in his stride. But, once that response was made, it is no sacrilege to find that language created to fulfill formal conventions displays statistical regularities conforming to those conventions.

Statistical Postscript

The above is an exercise in descriptive statistics. It is not clear that the application of inferential statistics would be very useful. For example, it is quite clear from Column (a) in Table I that the distribution of junctures at the various positions in the verse does not occur randomly. We could, if we wished, generate expected values on the basis of the null hypothesis that the number of junctures at each position is the result of a random distribution. On that basis, given 5202 internal junctures (the total of 6151 less the 949 at Position 0) and the 16 possible positions where juncture may occur, we arrive at an expected value of 325.1 junctures occurring at each position. If we then calculate the value of chi-squared for each position (at the extreme, chi-square for Position 16 is 311.3), we soon see that the values are so large that it is idle to complete the calculations and it is obvious that we are not dealing with a random distribution.

Nevertheless, it may be instructive to subject the major finding of this study to a brief inferential analysis. We hold on the basis of data displayed in Table V that the distribution of dactylic words ending with a consonant is not random. We test this as follows (cf. Table VII):

Table VII. Chi-square test for independence of word-ending and position

		Dactylic words		Totals
		with final consonant	with final vowel	
Ending	*d3*	93	34	127
at	*d12*	8	10	18
Position	15	108	229	337
	Totals	209	273	482

Chi-squared = 63.7020 with two degrees of freedom. $P < 0.0005$.

Once again, the value of chi-square is so great that there can be no question of a random distribution. This said, however, there remains the interesting question as to how or why this non-random distribution comes about, and this has been the object of our speculation.

Notes

1 OTT's observations are based upon the *Oxford Classical Text* of Mynors, 1969. The last verse is numbered 952, but 631 is a short verse, and there are no verses numbered 612 or 613. OTT's scheme for numbering positions is:
___1u2u3___4u5u6___7u8u9___10u11u12___13u14u15___16u.

2 STURTEVANT's views about the rest of the verse are less convincing, i.e., that there is an avoidance of such consonance in the rest of the verse. For *Aeneid* 12, OTT (1973: 26) reports consonance of ictus and accent as follows:

Foot 1	66.07%
Foot 2	31.51%
Foot 3	28.77%
Foot 4	37.09%

At least in the case of Foot 1, it seems *prima facie* unreasonable to speak of avoidance.

3 There are in fact 6174 words. We omit from our tally the twenty elided monosyllables and the three words in the short verse 631.

4 The observations in Table II are derived from texts largely supplied by the American Philological Associations's Repository of Greek and Latin Texts in Machine-Readable Form. This aid is gratefully acknowledged.

5 There is a problem of counting here. As Column (e) in Table I shows, the largest number of vocalic initials within the verse occurs at Position 10 (after the first syllable in Foot 4). Nevertheless, vocalic initials occur in only 31.08% of the junctures at Position 10, a bit less than the overall percentage of 32.68%. We believe that the large number of vocalic initials at Position 10 is a function of the large number of junctures of all types occurring at that position.

6 The matter with regard to elision has been explored by SOUBIRAN (1966) in Tableau A5, p. 518, and in the tables at pp. 496 ff. with attendant discussion.

7 So far as I can tell, SOUBIRAN (1966) nowhere notes this disparity between the numbers of elisions occurring at the end of spondees as opposed to dactyls.

8 Monosyllables of this sort are discussed by HELLEGOUARC'H (1964: 208 ff.). E.g., "... de plus ces mots doivent obligatoirement se trouver devant une initiale vocalique." (p. 208).

9 See the discussion of such monosyllables in HELLEGOUARC'H (1964: 167 ff.).

10 See comparable data in SOUBIRAN (1966: 592).

References

HELLEGOUARC'H, J., Le monosyllabe dans l'hexametre latin. Paris, Klincksieck 1964.

OTT, W., Metrische Analysen zu Vergil: Aeneis Buch XII. Tübingen, Niemeyer 1973.

SOUBIRAN, J., L'elision dans la poesie latine. Paris, Klincksieck 1966.

STURTEVANT, E.H., Word-ends and pauses in the hexameter. American Journal of Philology 42, 1921, 289-308.

Liaisons syllabiques dans les partages trochaïques de l'hexametre latin

J. Gérard, Paris

Abstract

Syllabic liaison in trochaic division in Latin hexameters.

The author sets out to define and apply to a wide series of samples an approach to the study of the liaisons existing between the final syllable of one word, in those cases in Latin hexameter poetry where the word division comes between the short syllables of a dactyl (trochaic division). His first step is to establish the conditions in which liaison normally occurs in poetry and prose. He then carries out a statistical examination of sixteen works, from Ennius to Juvenal, in order to give an opinion on the theory proposed by J.Soubiran that consonant -vowel liaison (and elision) are prefered to vowel-consonant liaisons when there occurs between two words a semantic break marked by punctuation. In such an investigation it is important to clearly distinguish the constituant elements of the vocabulary and to seperate words of full semantic value from link - words whose utilisation is different depending on whether there is a full stop or only a comma present. If for all forms of punctuation the theory seems to be confirmed, this would imply a special use of link-words beginning with a vowel in Latin hexameters, a supposition supported by a comparison of the utilisation of these link-words in poetry and in prose.

On a pu supposer avec raison que la division des brèves du dactyle entre une syllabe finale de mot et une syllabe initiale perturbe sensiblement le cours normal de l'hexamètre latin (PERRET 1954, 1955; GERARD 1980). Un intervalle entre deux mots provoque une pause, si minime soit-elle. Dans le cas du partage trochaïque, cette pause, faite au milieu d'un demi-pied, s'ajoute au temps requis par l'articulation des deux brèves et le biforme n'est plus isochrone du premier pied. Au lieu d'avoir l'équivalence habituelle: une longue égale deux brèves, nous avons: une longue égale une brève, plus une pause, plus une brève. La difficulté ainsi créée est d'autant plus grave qu'à l'intermot considéré intervient une pause de sens marquée par une ponctuation, forte ou faible.

On peut alors se demander si les poètes latins ne se sont pas résolus à admettre ces ponctuations sous certaines conditions de liaison syllabique entre les mots dont la finale et l'initiale fournissent les brèves du partage.

M.J.SOUBIRAN a montré que, dans l'hexamètre, l'élision ou la liaison consonne-voyelle (sigle CV) étaient préférées à la liaison voyelle-consonne (sigle VC), quand intervient une pause de sens entre les pieds II et III, III et IV, IV et V (ponctuation bucolique); les démonstrations qu'il en a données nous semblent tout à fait convaincantes (SOUBIRAN 1966 b, 1968, 1969) et nous faisons volontiers nôtre sa conclusion: la liaison syllabique assurée par l'élision ou par la séquence CV sert à pallier l'irrégularité fondamentale introduite par une pause de sens qui obligerait le récitant à marquer dans le débit un arrêt contraire à la bonne diction du vers.

Aussi, en cas de ponctuation trochaïque intervenant dans le partage des brèves, pourrait-on s'attendre à voir les poètes faire une différence entre des vers comme ceux-ci:

A Casta faue Lucina : tuus iam regnat Apollo
(VIRG. *Buc.* IV, 10)

B Qui Bauium non odit, amet tua carmina, Maeui.
(VIRG. *Buc.* III, 90)

Le premier présente une liaison VC (*Lucina tuus*, soit, dans la prononciation : *Lu-ci-na-tu-us)* qui distingue phonétiquement les mots et permet de faire coïncider coupe verbale et coupe syntaxique. Le second présente une liaison CV *(odit amet*, soit, dans la prononciation : *o-di-ta-met)* plus favorable à une diction liée.

Le problème est de savoir quels sont les vers les plus nombreux, ceux du type A ou ceux du type B.

PRINCIPES ET METHODES

Avec quelques nuances, nous emploierons la méthode utilisée par M.J. SOUBIRAN dans les recherches qu'il a consacrées aux liaisons syllabiques réalisées en cas d'intermot II/III, III/IV, IV/V (SOUBIRAN 1968:22 et 41-43).

1° Les ponctuations fortes et faibles sont, d'abord, comptées ensemble. La ponctuation forte est définie par les signes modernes usuels, point, point d'interrogation, d'exclamation, point et virgule, tirets, parenthèses, guillemets (GERARD 1980:16-21). La ponctuation faible est définie, à l'exclusion des cas précédents, comme une pause de sens intervenant à la fin d'une proposition à un mode personnel que suit une autre proposition également à un mode personnel.

2° Les liaisons syllabiques réalisées par l'élision et celles réalisées par la séquence CV sont considérées comme équivalentes.

3° Sont mises en comparaison, d'une part la fréquence des liaisons syllabiques CV et VC quand une ponctuation intervient à l'intermot considéré, d'autre part la fréquence de ces mêmes liaisons quand aucune ponctuation n'intervient à cette même place.

4° Dans les résultats obtenus, doivent ressortir les éléments constitutifs du vocabulaire déterminant la nature des liaisons syllabiques: d'un côté, les mots de sens plein, dont l'usage n'est soumis en principe à aucune contrainte, d'un autre côté, les outils de liaison, dont l'emploi est régi par des nécessités syntaxiques (subordonnants, coordonnants, adverbes ou pronoms).

Le fait métrique que nous nous proposons d'analyser est cependant assez différent de ceux qui ont retenu l'attention de M.J. SOUBIRAN, en raison même de sa situation, au milieu d'un pied. C'est pourquoi, aux principes ci-dessus définis, il est nécessaire d'adjoindre quelques précisions.

1° En cas de partage des brèves du dactyle, les liaisons syllabiques n'appartiennent pas toutes à l'un des types répertoriés: élision, CV, VC. Il peut y avoir hiatus. Le fait étant très rare, nous l'avons laissé en dehors de notre étude. Dans les textes que nous avons dépouillés, il n'y a, en effet, que 5 hiatus, dont 4 avec abrègement de la voyelle longue finale (LUCILIUS, 338, 1020; VIRG. *Buc.* II, 65; VIII, 108) et un de voyelle brève finale sur initiale vocalique brève (VIRG. *Buc.* II, 53).

De plus, il existe, dans l'ancienne poésie, des partages trochaïques dont la première brève est une finale fermée terminée en -s, cette consonne ne faisant pas position devant initiale con-

sonantique. Ce phénomène devient exceptionnel quand une ponctuation intervient entre les brèves: 2 exemples chez Ennius (428 vers), dans un même vers, avec une ponctuation trochaïque forte en III, une faible en V (vers 166, ed. J. VAHLEN); 2 exemples en V, dans les 605 vers de Lucilius retenus pour notre analyse (vers 92 et 181, ed. Fr. MARX), encore s'agit-il de fins de propositions incises que nous excluons de notre définition de la ponctuation faible (cf. infra). Nous laisserons donc aussi de côté ce type de liaison qui n'est admis, semble-t-il, qu'en l'absence de toute ponctuation.

2° Afin d'éliminer tout fait contestable, nous prendrons la définition de la ponctuation faible dans son sens le plus strict, en ne retenant que les pauses de sens intervenant à la fin d'une proposition à un mode personnel qui doit elle-même être suivie d'une autre proposition à un mode personnel.

En application de la définition restrictive énoncée ci-dessus on ne considérera pas qu'il y a pause de sens:

1) au début ou à la fin des propositions incises, qui ne se distinguent pas toujours nettement du contexte, comme dans les exemples suivants :

> At non ille, satum quo te mentiris, Achilles,
> talis in hoste fuit Priamo; ...
> (VIRG. *Aen.* II, 540 sq.)
>
> Ille quod exiguum restabat sanguinis Urbi
> hausit; ...
> (LUCAN. II, 140 sq.)
>
> omnis, ut est, igitur per se natura duabus
> constitit in rebus; ...
> (LUCR. I, 419 sq.)

2) avant ou après les verbes déclaratifs du type *ait*, *inquit*, ou leur équivalents comme *opinor*, *precor*, *reor*, *spero* mis en incise:

> Tandem urbanus ad hunc: "Quid te iuuat, inquit amice,
> (HOR. *Sat.* II, 6, 90)
>
> "Quis rubor iste poli? quod tam lugubre refulsit
> sidus?" ait; reddit trepido cui talia uirgo:
> (V.Fl. *Arg.* VIII, 58 sq.)
>
> ... Nati patrisque,
> alma, precor, miserere (potes namque omnia, nec te
> nequiquam lucis Hecate praefecit Auernis).
> (VIRG. *Aen.* VI, 116-118)

Utilisés à l'intérieur d'une proposition, ces verbes rentrent en effet dans la catégorie des incises; s'ils prennent place à la fin d'une phrase rédigée au style direct, on doit les considérer comme des enclitiques, la véritable pause de sens intervenant après eux.

3) entre les verbes et les conjonctions subordonnantes dans les locutions apparemment indissociables du type *fiet ut, adde quod, effice ut, accedit ut*:

fiet uti risu tremulo concussa cachinnent
(LUCR. I, 919)

Adde quod imbecilla nimis primordia fingit;
(LUCR. I, 847)

4) entre des verbes coordonnés du type *frangitque quatitque* :

quas animosi Euri assidue franguntque feruntque
(VIRG. *Georg.* II, 441)

Scintillant oculi dicisque facisque quod ipse
non sani esse hominis non sanus iuret Orestes.
(PERS. III, 117 sq.)

... neque enim uelocior ullas
itque reditque uias cognataue Tartara mauult.
(STAT. *Theb.* I, 101 sq.)

Cette double coordination ne permet pas de distinguer réellement deux propositions; elle représente une structure métrique commode et tout aussi fréquemment réalisée avec des substantifs qu'avec des verbes.

Seuls ont donc été retenus comme affectés d'une pause de sens faible, autrement dit d'une ponctuation faible, les intermots trochaïques marquant le partage:

A) entre des propositions juxtaposées que ne relie aucun mot subordonnant ou coordonnant:

Perdidit arma, locum uirtutis deseruit, qui
semper in augenda festinat et obruitur re
(HOR. *Epist.* I, 16,67 sq.)

B) entre des propositions principales et subordonnées:

- la subordonnée précédant la principale,

si laedit caupona, Ferentium ire uidebo;
(HOR. *Epist.* I, 17,8)

- la subordonnée suivant la principale,

In niue Lucana dormis ocreatus, ut aprum
cenem ego: tu piscis hiberno ex aequore uerris.
(HOR. *Sat.* II, 3,234 sq.)

C) entre des propositions coordonnées :

- le coordonnant ouvre la proposition,

non usquam prorepit et illis utitur ante
quaesitis sapiens, ...
(HOR. *Sat.* I, 1,37 sq.)

- le coordonnant est *-que* rattaché au premier mot,

Quod quoniam ratio reclamat uere negatque
credere posse animum, uictus fateare necessest.
(LUCR. I, 623 sq.)

Une rigoureuse application de ces principes exige un dépouillement très attentif des textes, ce qui nous a posé problème pour le choix d'un matériel d'étude. Dans notre thèse (GERARD 1980), l'extrême rareté du phénomène étudié, la ponctuation trochaïque forte, la relative facilité à la reconnaître grâce aux signes de ponctuation, nous avaient permis d'analyser la totalité des oeuvres retenues; il a donc été possible d'y rendre compte de l'intégralité des cas de ponctuation trochaïque forte. La détermination de la ponctuation trochaïque faible est beaucoup plus délicate, comme on vient de le voir, les faits étant beaucoup plus nombreux et divers. Deux solutions s'offraient alors : ou bien retenir quelques-unes seulement des oeuvres choisies pour l'étude de la ponctuation trochaïque forte et en faire l'analyse complète, ou bien reprendre toutes les oeuvres précédemment analysées en opérant cette fois sur des fragments de 1000 vers ou sur l'oeuvre entière quand elle n'atteint pas mille vers.

Nous nous sommes arrêté à cette seconde solution pour plusieurs raisons. Il nous a semblé intéressant de faire, à propos des liaisons syllabiques, une enquête parallèle à celle que nous avions déjà menée, d'autant plus qu'il était facile d'utiliser des relevés exhaustifs ayant porté sur l'intégralité des oeuvres en ce qui concerne les ponctuations trochaïques fortes. Il n'était pas non plus sans intérêt de conserver la perspective historique de notre premier travail. Il importait surtout de définir des méthodes d'enquête et de poser des hypothèses de travail avant d'entreprendre une analyse exhaustive du phénomène, compte tenu de l'extrême complexité des faits qui requièrent l'attention quand on avance dans l'analyse des éléments constitutifs du partage trochaïque marqué de ponctuation faible, en particulier quand on entreprend l'analyse verbale de ces éléments.

Notre lecteur voudra donc bien considérer ce travail comme exploratoire et méthodologique. Si, chemin faisant, nous croyons pouvoir ouvrir certaines perspectives à l'interprétation, il ne peut s'agir que d'hypothèses faites sous réserve des vérifications complémentaires que peuvent apporter l'analyse de la totalité des oeuvres (il en sera ainsi pour les liaisons syllabiques réalisées dans les partages trochaïques marqués de ponctuation forte) ou l'application de méthodes spécifiquement statistiques.

Il convient toutefois de rappeler ce que nous avons démontré à propos de la détermination des ponctuations trochaïques fortes (GERARD 1980:33-37): leur nombre ne dépend pas de la combinaison mathématique des éléments qui les constituent, partages trochaïques d'une part, ponctuations d'autre part. Les auteurs, qui manifestent une extrême réserve à en accepter l'emploi, s'y résolvent pour des raisons stylistiques et poétiques précises et rigoureuses, dans des proportions fort différentes d'une oeuvre à l'autre, même quand ils usent de manière égale des partages trochaïques et des ponctuations fortes. Le partage trochaïque marqué de ponctuation forte apparaît alors comme une figure métrique et stylistique dont l'usage relève de choix artistiques et n'obéit pas aux habitudes générales ordinaires de la langue et de la technique poétique. On peut en inférer qu'il y a quelque chance pour qu'il en soit de même lorsqu'une ponctuation faible intervient dans le partage des brèves et que nous nous trouvions en présence d'un fait très irrégulièrement réparti dans les oeuvres considérées.

Nos analyses ont donc porté sur les oeuvres suivantes, analysées dans leur entier quand il s'agit de ponctuations trochaïques fortes, sur tranches de 1000 vers quand il s'agit de ponctuations faibles, lues dans les éditions indiquées :

ENNIUS, *Ennianae poesis reliquiae* (ed. J.VAHLEN, Leibzig, Teubner 1928)

LUCILIUS, *C. Lucilii carminum reliquiae* (ed. Fr.MARX, Leibzig, Teubner 1904).

CATULLE, *Carmina*, pièce 64 (ed. G.LAFAYE, Paris, Coll.Univ.de France 1932).

LUCRECE, *De natura rerum* (ed. A.ERNOUT, Paris, Coll.Univ.de France 1941).

VIRGILE, *Les Bucoliques* (ed. E.de SAINT-DENIS, Paris 1942).
Les Géorgiques (ed. H.GOELZER, Paris 1926).
L'Enéide (I-VI, ed. H.GOELZER et A.BELLESSORT, Paris 1938; VII-XII, ed. R.DURAND et A.BELLESSORT, Paris 1941).

HORACE, *Les Epîtres* (ed. Fr.VILLENEUVE, Paris 1941).
Les Satires (ed. Fr.VILLENEUVE, Paris 1932).

OVIDE, *Les Métamorphoses* (ed. G.LAFAYE, Paris 1928).

LUCAIN, *Bellum ciuile* (ed. A.BOURGERY et M.PONCHONT, Paris 1926-1929).

STACE, *Les Silves* (ed. H.FRERE et H.J.IZAAC, Paris 1944).
La Thébaide (ed. A.KLOTZ, Leibzig, Teubner 1908).

PERSE, *Les Satires* (ed. A.CARTAULT, Paris 1929).

JUVENAL, *Les Satires* (ed. P.de LABRIOLLE et Fr.VILLENEUVE, Paris 1941).

VALERIUS FLACCUS, *Argonautica* (ed. J.H.MOZLEY, Loeb, Harvard Univ.Press 1963).

La totalité de ces oeuvres représente 68698 vers examinés.

Les tranches de mille vers ont été ainsi déterminées :

LUCRECE, I, 1-1000. VIRGILE, *Géorg.* II, IV, 1-459. *Aen.* II; VI, 1-196. HORACE, *Sat.* I, 1,2,3,8; II, 3,5,6. *Epist.* I, 1,7,13,15, 16,17,18; II, 3. OVIDE, *Métam.* I; IV, 1-221. LUCAIN, II; IV, 1-264. STACE, *Silu.* I, 1,2; II, 1,4; III, 1 (1-90), 2,5. *Theb.*I, IX, 1-280. JUVENAL, 1,2,3,4,5,6 (1-10). VALERIUS FLACCUS, *Arg.* II, VIII, 1-336.

Le matériel restreint, tranches de 1000 vers, plus oeuvres complètes quand elles n'atteignent pas mille vers, totalise 13920 vers (cf. infra, tableau I).

Dans ces 13 920 vers, nous avons relevé 191 partages trochaïques marqués d'une ponctuation forte et 428 marqués d'une ponctuation faible.

En cas de ponctuation faible, le détail de la répartition dans les catégories ci-dessus répertoriées en A), B), C) s'établit ainsi:

A) La ponctuation intervient entre des propositions juxtaposées que ne relie aucun mot subordonnant ou coordonnant: 32.

B) La ponctuation intervient entre des propositions principales et subordonnées:

- la subordonnée précède la principale: 44;

- la subordonnée suit la principale: 87 (dans 22 cas, le subordonnant n'est pas en tête de la subordonnée, il y a prolepse).

C) La ponctuation intervient entre des propositions coordonnées:

- le coordonnant ouvre la proposition: 234;
- le coordonnant est *-que* rattaché au premier mot de la proposition: 31.

On remarquera l'importance du nombre des propositions introduites par un mot de liaison: 65 subordonnées (87-22; cf. supra B), 234 coordonnées, au total 299 sur 428 partages marqués de ponctuation faible.

Le tableau I donne la répartition de ces ponctuations trochaïques faibles par pied, pour chaque oeuvre, ainsi que le pourcentage de fréquence par rapport au nombre de vers examinés.

Au vu de ces résultats chiffrés, on notera que la ponctuation trochaïque faible se rencontre, en moyenne générale calculée sur les 13 920 vers, dans la proportion de 3,1 %, mais que cette proportion varie de 1,2 % à 5,6 % quand on considère la répartition de détail dans les 16 oeuvres ou fragments examinés. Huit de ces oeuvres, soit la moitié, présentent un pourcentage compris entre 2,2 et 3 %, soit, dans l'ordre croissant, avec un écart de huit dixièmes entre les extrêmes :

2,2 %	*Bucoliques*
2,3 %	Lucain
2,5 %	*Silves*
2,7 %	Juvénal
2,8 %	*Métamorphoses*
3 %	*Géorgiques, Thébaïde, Argonautiques*

Les fréquences relevées dans les oeuvres de Lucilius (1,8 %) et de Catulle (2,0 %) sont assez proches de celles qui caractérisent le groupe précédent, à quelques dixièmes du pourcentage inférieur.

En revanche, nous trouvons des extrêmes : d'une part Ennius (1,2 %) et *L'Enéide* (1,4 %); de l'autre Lucrèce (4,1 %), Perse (4,9 %), Horace, *Epîtres* (5 %) et *Satires* (5,6 %) Dans ces quatre derniers cas, les pourcentages sont supérieurs de plus d'un point et même de plus de deux points aux 3 % des *Géorgiques*, de la *Thébaïde* et des *Argonautiques*, les trois oeuvres groupées de la série répertoriée ci-dessus pour lesquelles la fréquence d'usage

TABLEAU I. Fréquence des ponctuations trochaïques faibles.

Auteurs/ Oeuvres	Ponctuations trochaïques faibles							Nombre de vers
	I	II	III	IV	V	Total		
	n_i	n_i	n_i	n_i	n_i	n_i	%	
Ennius	1	1	1	1	1	5	1,2	428
Lucilius	1	4	5	-	1	11	1,8	605
Catulle	1	2	5	-	-	8	2,0	408
Lucrèce	9	8	5	1	18	41	4,1	1000
Buc.	6	1	6	1	4	18	2,2	829
Géorg.	9	2	4	-	15	30	3	1000
Enéide	8	-	3	1	2	14	1,4	1000
Hor. *Sat.*	10	3	10	2	31	56	5,6	1000
Hor. *Epist.*	10	4	5	6	25	50	5	1000
Métam.	16	1	-	-	11	28	2,8	1000
Lucain	14	1	1	-	7	23	2,3	1000
Perse	10	2	1	1	18	32	4,9	650
Silves	18	-	-	-	7	25	2,5	1000
Théb.	18	-	-	-	12	30	3	1000
Arg.	23	1	-	1	5	30	3	1000
Juvénal	11	-	3	-	13	27	2,7	1000
Total	165	30	49	14	170	428	3,1	13920

des ponctuations trochaïques faibles est la même que la fréquence moyenne calculée sur les 13920 vers.

Malgré cette apparente cohésion, la répartition des ponctuations trochaïques faibles à travers les oeuvres et fragments analysés n'est donc pas régulière, si l'on considère leur nombre et leur fréquence dans chaque oeuvre. Comme nous l'avions laissé prévoir ci-dessus, on pouvait s'attendre à une telle irrégularité à juger d'après les statistiques établies sur l'ensemble des ponctuations trochaïques fortes dans la totalité des oeuvres. La même irrégularité apparaît dans la distribution par pied. Toutefois, et en précisant encore qu'il s'agit ici de sondages sur des échantillons qui ne peuvent avoir que valeur indicative, les ponctuations trochaïques faibles apparaissent, en valeur absolue comme en valeur relative, plus nombreuses aux pieds I et V qu'aux autres pieds à partir de Lucrèce, dans 13 oeuvres sur 16; plus particulièrement, depuis les *Métamorphoses* jusqu'à Juvénal, elles n'apparaissent plus que très rarement aux pieds II, III et IV.

3° Les ponctuations trochaïques fortes sont, par nature, un cas limite des ponctuations faibles telles qu'elles viennent d'être définies, car elles interviennent non plus entre des propositions, mais entre des phrases. On en compte 191 dans les textes étudiés. Leur fréquence moyenne , dans l'ensemble des 13 920 vers, est de 1,4 %, à peu près la moitié de la fréquence moyenne des ponctuations trochaïques faibles. Leur répartition est encore plus irrégulière que dans le cas précédent. Catulle, dans la pièce 64, n'en a aucun exemple; sur les 15 autres textes analysés, 8 présentent des pourcentages de fréquence voisins de 1 %, soit, dans l'ordre croissant, avec un écart de cinq dixièmes:

0,7 %	*Argonautiques*
0,8 %	Lucilius, *Enéide*
1 %	*Métamorphoses, Silves, Thébaide*
1,2 %	Juvénal

Aux extrêmes, Lucrèce (0,3 %), Lucain (0,4 %), Ennius (0,5 %); d'autre part, *Géorgiques* (2 %), Horace, *Epîtres* (2,1 %), *Bucoliques* (3,3 %), Horace, *Satires* (4,2 %).

Le tableau II propose l'ensemble des résultats et permet de constater une plus grande irrégularité encore dans la répartition entre les pieds. L'extrême irrégularité de répartition apparaît nettement si l'on prête attention au fait que, par exemple, pour mille vers, il y a 8 ponctuations trochaïques fortes dans l'*Enéide*, 42 dans les *Satires* d'Horace, 4 seulement chez Lucain. Nous sommes en présence d'un phénomène tout à fait exceptionnel que nous avons étudié par ailleurs dans son intégralité (GERARD 1980). C'est l'une des raisons qui nous a poussé, pour l'analyse des liaisons syllabiques, à ne pas dissocier d'abord ponctuations fortes et ponctuations faibles, nous réservant d'ajouter une statistique particulière tenant compte alors de toutes les ponctuations trochaïques fortes relevées dans la totalité des oeuvres examinées. Il nous a semblé utile, en effet, de pouvoir saisir un état complet et indiscutable de ce fait de ponctuation exceptionnel, qui représente un cas limite, afin de comparer les résultats obtenus à ceux que fournit l'analyse fondée sur des extraits.

TABLEAU II. Fréquence des ponctuations trochaïques fortes.

Auteurs/ Oeuvres	Ponctuations trochaïques fortes							Nombre de vers
	I n_i	II n_i	III n_i	IV n_i	V n_i	Total n_i	%	
Ennius	-	1	1	-	-	2	0,5	428
Lucilius	-	-	2	2	1	5	0,8	605
Catulle	-	-	-	-	-	-	-	408
Lucrèce	-	-	2	-	1	3	0,3	1000
Buc.	1	3	17	-	6	27	3,3	829
Géorg.	4	4	6	-	6	20	2	1000
Enéide	-	2	4	-	2	8	0,8	1000
Hor. *Sat.*	-	2	16	1	23	42	4,2	1000
Hor. *Epist.*	4	2	5	1	9	21	2,1	1000
Métam.	4	1	1	-	4	10	1	1000
Lucain	-	-	-	-	4	4	0.4	1000
Perse	1	3	-	-	6	10	1,5	650
Silves	7	1	-	-	2	10	1	1000
Théb.	3	-	1	-	6	10	1	1000
Arg.	3	1	2	-	1	7	0,7	1000
Juvénal	1	-	-	-	11	12	1,2	1000
Total	28	20	57	4	82	191	1,4	13920

PROSODIE DES FINALES ET LIAISON SYLLABIQUE DANS LES PARTAGES TROCHAIQUES NON PONCTUES

La réalisation d'un partage trochaïque est soumise à des conditions si restrictives qu'il est difficile de prévoir quelle sera la liaison syllabique la plus fréquente par le seul examen de la prosodie des finales, comme M.J. SOUBIRAN a pu le faire en se plaçant d'un point de vue plus général (SOUBIRAN 1968:410-412). Etudiant les intermots II/III, III/IV, IV/V, il n'y avait pas, en effet, à se préoccuper de la quantité de la syllabe initiale, longue en début de pied par nature ou par position, mais seulement de la lettre initiale, voyelle ou consonne. Au contraire, dans le cas du partage trochaïque, syllabe finale et syllabe initiale doivent être et rester brèves, ce qui impose des restrictions particulières.

Ainsi, la finale qui donne la première brève, peut être de deux types:

1) syllabe brève fermée, du type *magnus;*

2) syllabe brève ouverte du type *magnă.*

Le partage n'est réalisable que si chacun de ces types est suivi respectivement d'une initiale vocalique brève (type 1) ou d'une initiale consonantique brève (type 2).

Se présente enfin une troisième éventualité: une élision intervient dans le partage. Elle ne peut se produire que sur initiale vocalique brève, la finale étant une voyelle brève ou longue (rarement une diphtongue) ou bien une syllabe brève terminée en -m. Là encore une condition supplémentaire rend difficile toute prévision : pour qu'il y ait encore partage trochaïque, la finale élidée doit être précédée d'une syllabe brève.

Ces trois conditions essentielles font intervenir des obligations de quantité concernant les syllabes et cela change les données du problème.

On peut en effet prévoir quelle sera la proportion des liaisons syllabiques permettant d'enchaîner le débit (élision et liaison CV) si l'on s'en tient aux seules possibilités de rencontre entre finales et initiales définies par une catégorie de lettres (consonne ou voyelle; exceptionnellement, à cause de la synalèphe, -m finale). M.J. Soubiran a pu faire le calcul suivant: quatre cinquièmes des finales et un tiers des initiales permettent de réaliser une liaison consonne-voyelle ou une élision (SOUBIRAN 1968:412). Mais il est inutile pour nous de pousser plus loin ce calcul statistique, si l'on considère le partage trochaïque. En effet, en ce cas, on notera que seulement deux types de finales (*magnŭs, magnă*) sur cinq (*magnōs, magnŭs, magnă, magnam, magnā*) se prêtent à la réalisation du partage; quant au nombre des initiales présentant voyelle ou consonne et pouvant compléter la figure métrique, il se trouve restreint puisque seules peuvent être retenues les syllabes brèves. Répertorier celles-ci en deux groupes (syllabes à initiale consonantique; syllabes brèves à initiale vocalique) reviendrait à dépouiller le vocabulaire latin en fonction, non seulement de la lettre initiale, mais encore de la quantité de la syllabe.

Devant cette difficulté, nous avons cru devoir adopter une méthode plus simple pour juger de la fréquence probable des types de liaisons syllabiques dans le partage trochaïque: la comparaison avec les liaisons comparables en prose.

Le problème, en effet, n'est pas tellement de prévoir quelle doit être statistiquement la liaison syllabique la plus fréquente en fonction des finales et des initiales de mots, mais de déterminer si les obligations métriques propres au partage trochaïque modifient quelque chose aux liaisons syllabiques ordinairement réalisées, c'est à dire, en général, imposées par le vocabulaire.

On laissera de côté le problème de l'élision qui ne peut être posé dans les mêmes termes en prose et en poésie. D'ailleurs, en cas de partage trochaïque non marqué de ponctuation, le nombre des élisions est faible: 307 sur 12582 partages (cf. tableau III), soit en moyenne 2,4 %. La proportion est un peu plus élevée quand une ponctuation, forte ou faible, intervient entre les deux brèves: 45 élisions sur 619 partages (cf. tableau IV), soit en moyenne 7,3 %.

La comparaison entre les liaisons syllabiques réalisées en prose et celles réalisées en poésie en cas de partage trochaïque non ponctué a donc été conduite de la façon suivante. Nous avons relevé le nombre de liaisons CV (consonne-voyelle) et VC (voyelle-consonne) réalisées en dehors de toute ponctuation :

1° dans la *Deuxième Philippique* de Cicéron;

2° dans les textes en hexamètres retenus pour l'étude des ponctuations trochaïques faibles. Rappelons qu'il s'agit des oeuvres complètes quand elles n'atteignent pas mille vers et, pour les autres, de tranches de mille vers dont les références sont données à la page 175, soit en tout 13920 hexamètres.

1° *Deuxième Philippique :*

CV	1517	37,1 %
VC	2575	62,9 %
	4092	

2° Textes poétiques :

CV	2961	24,1 %
VC	9313	75,9 %
	12274	

A en juger d'après les échantillonages analysés, les liaisons VC représentent donc en prose 62,9 % des ponctuations examinées, en poésie 75,9 %; les liaisons CV respectivement 37,1 % et 24,1 %.

Si l'on remarque qu'en prose la quantité des syllabes finales ou initiales est indifféremment brève ou longue, alors qu'elle est obligatoirement brève dans le partage trochaïque, on peut émettre l'hypothèse que cette dernière restriction a favorisé la liaison VC au détriment de la liaison CV. Il y a d'ailleurs des raisons pour que la liaison CV soit devenue, dans ce cas, plus difficilement réalisable. En effet, les finales brèves fermées représentent 25 % du vocabulaire, mais ne peuvent rester telles que si l'initiale du mot suivant est à la fois vocalique et brève; or, quelle que soit la quantité de la syllabe, il n'y a environ qu'une chance sur trois pour que l'initiale soit vocalique (SOUBIRAN 1966a:591; 1968:411); s'il faut de plus que cette initiale soit brève, cela réduit encore le nombre des chances de rencontre CV.

Mais nous avons comparé une oeuvre de prose à un ensemble d'oeuvres poétiques. Il convient de donner le détail des résultats; le tableau III les présente par auteur et par oeuvre. On a ajouté une colonne concernant les élisions, afin de voir si ce fait, exclu de l'enquête générale, est capable d'en modifier sensiblement l'orientation. Cette statistique servira ultérieurement à établir une comparaison entre les fréquences des liaisons CV/VC réalisées en cas de partage trochaïque marqué d'une ponctuation et celles des mêmes liaisons réalisées en cas de partage trochaïque non marqué de ponctuation. Pour le calcul des pourcentages, les cas d'élision ont été comptés avec les liaisons CV puisque, dans les deux cas, le phénomène phonétique assure un meilleur enchaînement du débit (cf. Principes et méthodes, 2°).

Les résultats détaillés fournis par le tableau III appellent quelques remarques indicatives, sous la réserve déjà faite qu'il s'agit d'hypothèses à partir de relevés partiels.

Il est d'abord intéressant d'ordonner les fréquences d'emploi des liaisons syllabiques en distinguant d'une part El + CV, d'autre part VC.

	EL+CV %	VC %
Lucain	19	81
Ennius	20	80
Catulle; *Argonautiques*	21	79
Enéide; Thébaide	22	78
Silves	23	77
Géorgiques	25	75
Lucrèce; *Métamorphoses;* Perse	27	73
Juvénal	30	70
Bucoliques; Horace, *Epîtres*	33	67
Horace, *Satires*	34	66
Lucilius	38	62

TABLEAU III. Liaisons syllabiques réalisées en cas de partage non marqué de ponctuation.

Auteurs/ Oeuvres	Eli-sion n_i	CV n_i	VC n_i	Total n_i	El+CV %	VC %	Nombre de vers n_i
Ennius	4	64	268	336	20	80	428
Lucilius	49	95	232	376	38	62	605
Catulle	12	40	200	252	21	79	408
Lucrèce	50	163	574	787	27	73	1000
Buc.	33	221	505	759	33	67	829
Géorg.	16	192	628	836	25	75	1000
Enéide	21	196	752	969	22	78	1000
Hor. *Sat.*	25	229	483	737	34	66	1000
Hor. *Epist.*	8	293	621	922	33	67	1000
Métam.	20	261	786	1070	27	73	1000
Lucain	7	192	865	1064	19	81	1000
Perse	14	106	321	441	27	73	650
Silves	15	234	832	1081	23	77	1000
Théb.	13	223	827	1063	22	78	1000
Arg.	17	215	868	1100	21	79	1000
Juvénal	3	234	552	789	30	70	1000
Total	307	2961	9314	12582	26	74	13920

Note: Pour les oeuvres d'Ennius et de Lucilius, on n'a pas tenu compte des partages trochaïques dans lesquels *-s* ne fait pas position. Soit, chez Ennius, 37 cas, ainsi répartis entre les pieds: I, 3 cas; II, 4 cas; III, 7 cas; IV, 21 cas. Chez Lucilius, 77 cas: I, 5 cas; II, 14 cas; III, 9 cas; IV, 3 cas; V, 46 cas. Sont également exclus du relevé deux cas de hiatus de brève sur brève: Lucilius, vers 358 et 1020.

On note que pour 4 oeuvres sur 16, les fréquences d'emploi des liaisons syllabiques se tiennent entre 33 et 38 % d'une part (El + CV), 62 et 67 % d'autre part (VC), fréquences très voisines de celles que nous avons relevées en prose dans la *Deuxième Philippique*, respectivement 37,1 % (CV) et 62,9 % (VC). Pour les 12 autres oeuvres, ces fréquences se tiennent entre 20 et 30 % (liaisons El + CV), 70 et 80 % (liaisons VC), soit des pourcentages légèrement inférieurs ou légèrement supérieurs aux pourcentages correspondants relevés en prose.

Quelle est alors la situation quand une ponctuation intervient entre les brèves du partage?

LIAISON SYLLABIQUE DANS LES PARTAGES TROCHAIQUES PONCTUES

Les textes ou fragments de textes analysés comprennent au total 619 partages trochaïques ponctués: 428 ponctuations faibles; 191 ponctuations fortes.

Au total, les liaisons syllabiques se répartissent ainsi:

EL + CV	45 + 313	58 %
VC	261	42 %

Le tableau IV donne le détail, par auteur et par oeuvre. Les cas d'élision sont comptés dans les chiffres des colonnes CV, mais leur nombre est porté entre parenthèses à côté du chiffre global correspondant.

La ponctuation paraît avoir modifié totalement le rapport des liaisons syllabiques dans le sens indiqué par M. J. Soubiran. Si nous considérons les fréquences moyennes calculées sur l'ensemble des vers analysés, soit CV = 58 %, VC = 42 %, nous constatons que les liaisons CV sont plus nombreuses que les liaisons VC. La tendance était inverse quand il s'agissait des partages trochaïques non ponctués et de la prose. Mais cette comparaison ne peut être donnée qu'à titre indicatif, le fait de ponctuation trochaïque n'étant pas homogène pour tous les auteurs. Il est toutefois intéressant de classer par ordre les fréquences de ces liaisons syllabiques comme nous l'avons fait ci-dessus.

	EL + CV %	VC %
Catulle	0	100
Juvénal	49	51
Perse; *Enéide*	50	50
Bucoliques	53	47
Lucrèce	55	45
Lucilius	56	44
Horace, *Satires*	57	43
Géorgiques; Horace, *Epîtres*	58	42
Silves	63	37
Lucain	67	33
Métamorphoses; Thébaide	68	32
Argonautiques	70	30
Ennius	71	29

TABLEAU IV. Liaisons syllabiques réalisées en cas de partage trochaïque marqué de ponctuation

Auteurs/ Oeuvres	Ponctuations fortes		Ponctuations faibles		Total				Nombre de vers
	CV n_i	VC n_i	CV n_i	VC n_i	CV n_i	VC n_i	CV %	VC %	n_i
Ennius	2	–	3	2	5	2	71	29	428
Lucilius	4 (3)	1	5	6	9	7	56	44	605
Catulle	–	–	–	8	–	8	0	100	408
Lucrèce	1	2	23 (2)	18	24	20	55	45	1000
Buc.	10	17	14 (4)	4	24	21	53	47	829
Géorg.	9	11	20 (2)	10	29	21	58	42	1000
Enéide	1	7	10 (4)	4	11	11	50	50	1000
Hor. *Sat.*	20 (1)	22	36 (1)	20	56	42	57	43	1000
Hor. *Epist.*	9	12	32 (3)	18	41	30	58	42	1000
Métam.	6	4	20 (1)	8	26	12	68	32	1000
Lucain	–	4	18 (1)	5	18	9	67	33	1000
Perse	2	8	19 (2)	13	21	21	50	50	650
Silves	4 (3)	6	18 (6)	7	22	13	63	37	1000
Théb.	6	4	21 (5)	9	27	13	68	32	1000
Arg.	2	5	24 (4)	6	26	11	70	30	1000
Juvénal	4	8	15 (3)	12	19	20	49	51	1000
Total	80 (7)	111	278(38)	150	358	261	58	42	13920

Notes: 1) Les chiffres entre parenthèses indiquent le nombre d'élisions comptées dans le nombre des liaisons CV.

2) Il n'a pas été tenu compte des cas où *-s* ne fait pas position, soit deux cas, dans un même vers d'Ennius, l'un de ponctuation forte, l'autre de ponctuation faible:

Inicit inritatus: tenet occasus, iuuat res.
(166, ed. J.VAHLEN)

On constate que, mis à part le cas exceptionnel de Catulle dont le texte analysé (pièce 64) n'offre aucun exemple de liaison CV, la fréquence des liaisons CV est plus grande que celle des liaisons VC dans 12 textes sur 15; dans les 3 autres, ces deux types de liaison sont à peu près à égalité. Voilà vérifiée dans le détail (et toujours sous réserve qu'il s'agit d'extraits) la tendance qu'indiquaient les fréquences moyennes de 58 % et 42 %. D'où la nécessité de poser une nouvelle hypothèse: si, dans le cas des partages trochaïques ponctués, on voit s'inverser les fréquences d'emploi des liaisons CV/VC par rapport aux liaisons réalisées dans les partages trochaïques non ponctués, doit-on penser que la simple intervention d'une pause marquée entre les brèves du dactyle incite le poète à user plus volontiers de la liaison CV ? A cela, il y aurait une raison phonétique qui est à la base même de la théorie de M.J. Soubiran: la liaison CV est plus apte que la liaison VC à réaliser l'enchaînement entre deux mots que sépare une pause de sens, en un point de l'hexamètre où la division des brèves crée déjà une certaine difficulté rythmique. On pourrait toutefois penser également que le vocabulaire utilisé en cas de ponctuation trochaïque est responsable d'un tel état de choses si, par exemple, les mots qui fournissent la deuxième brève du partage sont plus nombreux à avoir une initiale vocalique (= liaison CV) qu'une initiale consonantique (= liaison VC), ce qui pourrait être le cas des mots introducteurs de propositions, subordonnants ou coordonnants. Il convient donc d'analyser la structure verbale des partages trochaïques marqués de ponctuation.

STRUCTURE VERBALE ET LIAISON SYLLABIQUE DANS LE PARTAGE TROCHAIQUE PONCTUE

L'étude de la structure verbale des partages trochaïques marqués de ponctuation doit en premier lieu prendre en considération les mots dont l'initiale fournit la deuxième brève du partage. En effet, les pauses de sens, fortes ou faibles, interviennent devant ces mots, à la fin d'un membre syntaxique (proposition à un mode personnel) et devant un autre membre syntaxique (proposition à un mode personnel). On doit donc s'attendre à trouver à cette place, par définition, nombre d'outils de liaison, coordonnants ou subordonnants: la structure syntaxique en impose l'usage en même temps qu'elle définit la pause. Parmi eux, il faut prévoir l'emploi de *nec, neque, sed, nisi, quod, quis, quibus*, qui réalisent la liaison VC (initiale consonantique), mais aussi l'emploi de *at, et, an, ut, ubi, utrum*, qui réalisent la liaison CV (initiale vocalique). Peuvent aussi se rencontrer à cette place des mots de sens plein à initiale brève, venant après la pause de sens, soit que les propositions se trouvent juxtaposées, soit qu'une prolepse amène l'un d'eux en tête du membre syntaxique, le subordonnant étant alors inclus dans la proposition.

Ainsi se définissent deux catégories, sur un plan très général; nous en préciserons ultérieurement le contenu. Mais on reconnaît dès maintenant que si rien ne peut laisser prévoir l'usage d'un mot de sens plein en tête d'un membre syntaxique, les outils de liaison, eux, y venant naturellement, peuvent représenter une contrainte lexicale dont le rôle peut être déterminant dans la réalisation des liaisons syllabiques.

Le problème à résoudre prend donc deux aspects:

1° Les liaisons CV sont-elles plus nombreuses que les liaisons VC dans chaque de ces deux catégories?

2° Si elles ne sont plus nombreuses que dans la catégorie des mots de liaison, ce fait est-il dû à une plus grande fréquence des mots de liaison à initiale vocalique dans l'usage général de la langue?

La réponse à ces deux questions doit permettre d'apprécier les contraintes imposées par le vocabulaire dans le partage trochaïque ponctué et le rôle qu'il joue dans la réalisation des liaisons syllabiques.

I. - Définition des catégories de mots dont l'initiale fournit la deuxième brève des partages trochaïques ponctués.

1° Les mots de sens plein.

La définition générale de ces mots va de soi: *amator, auis, diem, inultus, domus, recusat,* etc... en donnent exemple.

Nous avons toutefois adopté certaines conventions:

a) Sont considérés comme mots de sens plein les formes d'auxiliaires du type *habet, erit;* l'adverbe *satis;* les formes des adjectifs *meus, tuus, suus.* Placés en tête d'un membre syntaxique, ces mots prennent généralement un sens fort qui, à notre avis, permet de les classer ainsi. Leur emploi, en cette position, ne relève pas d'une contrainte lexicale, mais d'un choix stylistique. Pour exemples:

Huc huc aduentate, meas audite querelas,
(CATULLE, LXIV, 195)

At si condoluit temptatum frigore corpus
aut alius casus lecto te adflixit, habes qui
adsideat, fomenta paret, medicum roget,
(HOR. *Sat.* I, 1, 80-82)

b) Sont relevés parmi les mots de sens plein ceux d'entre eux auxquels s'attache le coordonnant *-que,* car, pour la réalisation de la liaison syllabique, ce coordonnant ne compte pas plus que le subordonnant quand il y a prolepse.

Pour qu'il n'y ait cependant aucune ambiguité, nous avons classé ces mots sous le sigle A.

c) Nous distinguerons aussi un groupe B, comprenant les formes pronominales des types *eum, ego, tibi, sibi,* le pronom *nihil* et les interjections *io, papae, nefas.*

Ces mots ne sont pas des outils de liaison, ils n'en jouent pas le rôle syntaxique. On pourrait toutefois considérer qu'ils constituent une contrainte lexicale, bien que certains emplois des formes pronominales relèvent souvent, en tête de proposition d'un effet stylistique. Nous les avons donc classés dans la catégorie des mots de sens plein, mais à part, pour ne créer aucune confusion dans l'interprétation des résultats statistiques.

2° Les mots outils.

Cette expression sera employée de préférence à celle de "outils de liaison". En effet, nous avons classé dans cette catégorie tous les mots dont l'usage constitue réellement une contrainte lexicale: les coordonnants (autres que *-que*), les subordonnants, les adverbes qui par leur sens assurent ordinairement une coordination (*dehinc, ibi, ita, item, modo, simul, tamen, ubique, uelut*), quelques interrogatifs, et les prépositions.

En dehors des adverbes déjà cités, ces mots sont en nombre limité, car ils doivent tous présenter une initiale brève. Ce sont:

- pour ceux qui commencent par une voyelle: *an, at, et, ubi, ut, uti,* les formes de *uter; ab, ad, in, ob.*
- pour ceux qui commencent par une consonne: *nec, neque, nisi, quibus, quid, quis, quod, quot, sed, uel; per.* Un cas exceptionnel: *si* s'élide une fois sur initiale vocalique brève (Lucilius, 268).

II. - Répartition des liaisons CV/VC entre les catégories de vocabulaire: mots de sens plein / mots-outils.

1° Ensemble des partages trochaïques ponctués.

Nous conviendrons de désigner ainsi les partages trochaïques marqués de ponctuation faible et de ponctuation forte, considérés ensemble, étant rappelé que cet ensemble ne représente que les partages trochaïques ponctués relevés dans les extraits auxquels nous nous sommes arrêté. Les nombres et pourcentages proposés ne représentent que des analyses fragmentaires et les remarques que nous en tirons sont faites à titre indicatif.

Si l'on tient compte de toutes les ponctuations, fortes et faibles, ainsi relevées, on constate que les liaisons VC sont nettement plus nombreuses que les liaisons CV dans la catégorie des mots de sens plein, les liaisons CV étant en revanche plus nombreuses que les liaisons VC dans la catégorie des mots-outils. Soit pour l'ensemble des textes analysés:

	CV		VC	
	n_i	%	n_i	%
Mots de sens plein	69	30	164	70
Mots-outils	289	75	97	25

Le tableau V donne le détail de nos relevés par oeuvre avec la distinction des groupes A et B dans la catégorie des mots de sens plein (cf. supra, p. 188).

TABLEAU V. Répartition des liaisons syllabiques entre mots de sens plein et mots-outils dans l'ensemble des partages trochaïques marqués de ponctuation.

Auteurs/ Oeuvres	Mots de sens plein						Mots outils		Total général		Nombre de vers
	Groupe A		Groupe B		Total						
	CV	VC	CV	VC	CV	VC	CV	VC	CV	VC	
Ennius	-	2	-	-	-	2	5	-	5	2	428
Lucilius	4	4	-	-	4	4	5	3	9	7	605
Catulle	-	2	-	3	-	5	-	3	-	8	408
Lucrèce	4	10	1	0	5	10	19	10	24	20	1000
Buc.	8	15	3	1	11	16	13	5	24	21	829
Géorg.	3	9	3	2	6	11	23	10	29	21	1000
Enéide	1	10	-	-	1	10	10	1	11	11	1000
Hor. *Sat.*	17	27	3	1	20	28	36	14	56	42	1000
Hor. *Epist.*	10	21	-	1	10	22	31	8	41	30	1000
Métam.	-	8	-	-	-	8	26	4	26	12	1000
Lucain	1	6	-	-	1	6	17	3	18	9	1000
Perse	2	12	-	1	2	13	19	8	21	21	650
Silves	6	8	-	-	6	8	16	5	22	13	1000
Théb.	-	5	-	-	-	5	27	8	27	13	1000
Arg.	1	5	-	2	1	7	25	4	26	11	1000
Juvénal	2	9	-	-	2	9	17	11	19	20	1000
Total	59	153	10	11	69	164	289	97	358	261	13920

La lecture du tableau V appelle les remarques suivantes:

1° Les liaisons VC sont plus nombreuses que les liaisons CV, exception faite du texte de Lucilius, dans la catégorie des mots de sens plein.

2° Les liaisons CV sont plus nombreuses que les liaisons VC, sans exception, dans la catégorie des mots-outils.

3° Si, dans le tableau IV, les liaisons CV apparaissent plus nombreuses que les liaisons VC, exceptions faites des textes de Catulle, de l'*Enéide*, de Perse et de Juvénal, le fait est dû au nombre de liaisons CV réalisées dans la catégorie des mots-outils, polysyllabes à initiale brève ou monosyllabes brefs, donc à l'usage d'un vocabulaire requis par l'articulation des propositions dans la phrase. On peut supposer qu'il y a, dans ces conditions, contrainte lexicale et non pas, comme le suppose M.J. SOUBIRAN à propos des ponctuations intervenant entre les pieds, préférence poétique pour ce type de liaison qui assurerait un meilleur enchaînement du rythme de l'hexamètre. Cela requiert attention et enquête plus minutieuse, en premier lieu, de préciser quelle est la situation dans chaque catégorie de partages trochaïques, ceux que marque une ponctuation forte, ceux que marque une ponctuation faible.

2° Partages trochaïques marqués d'une ponctuation forte.

On se rappelle que dans cette catégorie de partages, les liaisons VC sont plus nombreuses que les liaisons CV dans 11 textes sur 16 (cf. tableau IV). Il importe de savoir ce qu'il en est, dans ce cas, de l'usage des mots-outils et de celui des mots de sens plein.

Comme nous l'avons signalé ci-dessus, nous avons à notre disposition des relevés exhaustifs tenant compte de tous les cas de ponctuation trochaïque forte dans la totalité des oeuvres retenues pour cette étude (GERARD 1980); nous les utiliserons donc de préférence aux relevés partiels portant sur des tranches de 1000 vers, auxquels, d'ailleurs, nous pourrons constater qu'ils apportent confirmation. Dans ces conditions, l'analyse concerne 68698 vers dans lesquels se rencontrent 651 partages trochaïques marqués de ponctuation forte.

Les liaisons syllabiques sont ainsi réparties, dans l'ensemble:

	CV		VC	
	n_i	%	n_i	%
Mots de sens plein	108	32	230	68
Mots-outils	188	60	125	40
	296	45	355	55

Si, d'autre part, on veut bien se reporter au tableau IV, qui donne la répartition des liaisons syllabiques en cas de partages trochaïques marqués de ponctuation forte, et la comparer avec la colonne "total général" du tableau VI, qui donne les mêmes indications, mais cette fois, non plus pour des tranches de 1000 vers, mais pour la totalité des oeuvres, on constatera que la tendance des résultats est inchangée. Dans 10 oeuvres sur 16, les liaisons VC sont plus nombreuses que les liaisons CV, et ce sont les mêmes d'une statistique à l'autre; la seule modification concerne Lucain qui, dans les 1000 vers analysés, présentait 4 ponctuations trochaïques fortes en tout, réalisant une liaison VC, alors que dans toute l'oeuvre, il s'en rencontre 22, dont 10 réalisent la liaison VC.

La lecture du tableau VI appelle les remarques suivantes:

1° Les liaisons VC sont plus nombreuses que les liaisons CV dans la catégorie des mots de sens plein dans 13 oeuvres sur 16; dans les 3 textes restants, 2 n'ont aucun exemple de ponctuation trochaïque forte; il n'est que chez Lucilius que nous relevons 2 liaisons CV pour 1 liaison VC.

2° Dans la catégorie des mots-outils, sur 16 textes, les liaisons CV sont plus nombreuses que les liaisons VC dans 8 textes, les liaisons VC plus nombreuses que les liaisons CV dans 5; elles sont à égalité dans 2 textes; il n'y a aucun partage trochaïque marqué de ponctuation forte dans la pièce 64 de Catulle.

On émettra donc l'hypothèse que les poètes ne recherchent pas une liaison CV quand intervient une ponctuation forte entre les brèves du partage trochaïque et que l'usage des mots-outils à initiale vocalique ou à initiale consonantique reste indifférent.

Cela met en lumière que, si les liaisons CV sont plus nombreuses

que les liaisons VC dans la statistique générale qui prenait en considération les partages trochaïques marqués de ponctuation forte et de ponctuation faible, cela est bien dû à l'emploi des mots-outils dans la catégorie des partages trochaïques marqués de ponctuation faible, ce qu'il convient d'étudier dans une analyse particulière.

TABLEAU VI. Répartition des liaisons syllabiques entre mots de sens plein et mots-outils dans les partages trochaïques marqués de ponctuation forte (totalité des oeuvres: 68698 vers).

Auteurs/ Oeuvres	Mots de sens pleins						Mots outils		Total général	
	Groupe A		Groupe B		Total					
	CV	VC	CV	VC	CV	VC	CV	VC	CV	VC
Ennius[1]	-	-	-	-	-	-	2	-	2	-
Lucilius	2	1	-	-	2	1	2	-	4	1
Catulle	-	-	-	-	-	-	-	-	-	-
Lucrèce	-	2	1	-	1	2	7	7	8	9
Buc.[2]	6	14	2	-	8	14	2	3	10	17
Géorg.	5	9	2	1	7	10	8	11	15	21
Enéide	4	25	3	3	7	28	10	9	17	37
Hor. *Sat.*	13	32	4	1	17	33	25	12	42	45
Hor. *Epist.*	7	14	2	-	9	14	5	6	14	20
Métam.	9	30	4	-	13	30	53	25	66	55
Lucain	4	6	-	1	4	7	8	3	12	10
Perse	-	6	1	1	1	7	1	1	2	8
Silves	4	10	1	1	5	11	3	4	8	15
Théb.	25	36	5	5	30	41	48	30	78	71
Arg.	1	15	1	7	2	22	7	3	9	25
Juvénal	2	10	-	-	2	10	7	11	9	21
Total	82	210	26	20	108	230	188	125	296	355

Notes: 1) La liaison *inritatu(s) : tenet* (Ennius, 166 VAHLEN) n'est pas comptée (cf. tableau IV, n.3).
2) *Buc.* II, 53 présente un hiatus de brève sur brève; cette liaison n'a pas été prise en considération.
Pour la division en groupes A et B, cf. supra, p. 188

3° Partages trochaïques marqués de ponctuation faible.

Rappelons d'abord les résultats de la statistique générale: sur 619 partages trochaïques ponctués, 358 présentent une liaison CV, dont 69 sur initiale vocalique de mots de sens plein, 289 sur initiale vocalique de mots-outils.

Si l'on ne considère que les partages trochaïques marqués de ponctuation faible, on trouve:

Pour les chiffres d'ensemble:

	CV n_i	CV %	VC n_i	VC %
Mots de sens plein	30	25	90	75
Mots-outils	248	81	60	19
	278	65	150	35

Le tableau VII donne le détail des relevés.

La lecture de ce tableau appelle les remarques suivantes:

1° Catégorie des mots de sens plein: les liaisons VC sont plus nombreuses que les liaisons CV dans 14 oeuvres sur 16; les deux types de liaison syllabique sont à égalité dans les *Silves;* on relève, dans les *Bucoliques*, 3 liaisons CV pour 2 liaisons VC. Il convient de noter que les liaisons CV deviennent très rares dans les oeuvres classées, dans la liste chronologique, après les *Métamorphoses* (5 liaisons CV en 7000 vers), alors que dans les 9 premières oeuvres de la liste il s'en trouve 25, en 6920 vers. La situation apparaît inverse dans la catégorie des liaisons VC, attestées dans toutes les oeuvres. Nous répétons qu'il s'agit ici de constatations valables pour les fragments analysés.

2° Catégorie des mots-outils: les liaisons CV sont considérablement plus nombreuses que les liaisons VC dans 14 oeuvres sur 16; les deux types de liaison sont à égalité chez Lucilius; dans la pièce 64 de Catulle, il n'y a aucune liaison CV et 3 liaisons VC.

3° Total général: le fait que les liaisons CV soient, dans la catégorie des mots-outils, de 2 à 10 fois plus nombreuses, suivant les oeuvres, que les liaisons VC, a pour conséquence que, les deux catégories, mots de sens plein et mots-outils, se

trouvant réunies, ces mêmes liaisons CV restent plus nombreuses que les liaisons VC. Cette situation est bien le fait du nombre élevé de mots-outils à initiale vocalique utilisés, réalisant la liaison CV.

TABLEAU VII. Répartition des liaisons syllabiques entre mots de sens plein et mots-outils dans les partages trochaïques marqués de ponctuation faible.

Auteurs/ Oeuvres	Mots de sens plein: Groupe A CV	Groupe A VC	Groupe B CV	Groupe B VC	Total CV	Total VC	Mots outils CV	Mots outils VC	Total général CV	Total général VC
Ennius[1]	-	2	-	-	-	2	3	-	3	2
Lucilius	2	3	-	-	2	3	3	3	5	6
Catulle	-	2	-	3	-	5	-	3	-	8
Lucrèce	4	10	1	-	5	10	18	8	23	18
Buc.	2	1	1	1	3	2	11	2	14	4
Géorg.	-	4	1	1	1	5	19	5	20	10
Enéide	-	4	-	-	-	4	10	-	10	4
Hor.*Sat.*	9	12	1	1	10	13	26	7	36	20
Hor.*Epist.*	4	13	-	1	4	14	28	4	32	18
Métam.	-	6	-	-	-	6	20	2	20	8
Lucain	1	5	-	-	1	5	17	-	18	5
Perse	1	6	-	-	1	6	18	7	19	13
Silves	3	3	-	-	3	3	15	4	18	7
Théb.	-	4	-	-	-	4	21	5	21	9
Arg.	-	2	-	-	-	2	24	4	24	6
Juvénal	-	6	-	-	-	6	15	6	15	12
Total	26	83	4	7	30	90	248	60	278	150

Note: La liaison *occasu(s), iuuat res* (Ennius, 166 VAHLEN) n'est pas comprise dans ce relevé (cf. tableau IV, n.3).

Voilà donc à première vue une situation paradoxale. Nous étions parti d'une hypothèse faite par M.J. SOUBIRAN: la liaison CV est préférée à la liaison VC quand, à l'intermot, intervient une ponctuation; cela serait dû au fait qu'elle pallie la gêne causée par cette ponctuation dans le débit normal de l'hexamètre. Or il se révèle que, si les pourcentages calculés en tenant compte de toutes les ponctuations, fortes et faibles, semblent confirmer l'hypothèse, celle-ci cesse de se vérifier si l'on ne prend en considération que le cas extrême des ponctuations fortes. Dans cet ordre, les liaisons VC sont plus nombreuses que les liaisons CV, en raison même du nombre élevé des mots de sens plein à initiale consonantique fournissant la deuxième brève du partage; l'emploi de ces mots, après ponctuation forte, échappe à toute contrainte lexicale ou syntaxique.

L'hypothèse paraît se vérifier si, au contraire, on prend en considération les seuls partages trochaïques marqués de ponctuation faible: les liaisons CV sont alors considérablement plus nombreuses que les liaisons VC dans la presque totalité des oeuvres examinées. Mais une analyse plus attentive de la structure verbale nous montre que dans ce cas le grand nombre de liaisons CV s'explique par l'emploi fréquent de mots-outils à initiale vocalique qui représentent une contrainte lexicale puisqu'ils ont pour fonction d'introduire une proposition après ponctuation faible. On peut douter dans ces conditions qu'il s'agisse d'un fait de phonétique concernant le débit du vers et se demander si ce n'est pas un simple fait de vocabulaire. Il convient donc d'examiner si l'usage exceptionnellement fréquent de mots-outils à initiale vocalique que nous venons de constater (sur 5 liaisons assurées par des mots -outils, dans les conditions définies, 4 sont réalisées par des mots-outils à initiale vocalique) correspond à un usage courant d'une part, dans l'hexamètre, quand il y a ponctuation faible en dehors du partage trochaïque, d'autre part, dans la prose.

C'est essentiellement dans cette partie de l'enquête qu'il était nécessaire de procéder par sondages: les faits que nous allons devoir analyser sont extrêmement divers et nombreux; il n'était pas envisageable de chercher à faire d'emblée des relevés exhaustifs avant de savoir si pouvaient se dégager des tendances susceptibles d'apporter quelques éclaircissements sur la question. Aussi, et nous le répétons, les relevés statistiques qui vont suivre et les remarques qu'ils suggèrent, doivent-ils être pris à titre indicatif, non comme des conclusions définitives. Mais nous estimons qu'en procédant ainsi, non seulement nous pouvons

ouvrir la voie à des méthodes originales d'analyse verbale, mais encore que les hypothèses suggérées par ces analyses sont loin d'être gratuites ou mal fondées. Nous signalerons, chemin faisant, les raisons qui nous incitent à leur reconnaître quelque valeur, ne fût-ce que par la constance et l'homogénéité des résultats qui s'en dégagent. Nous suivrons l'ordre suivant.

Dans un premier temps, nous analyserons dans le détail la constitution du vocabulaire dans la catégorie des mots-outils dans les partages trochaïques marqués de ponctuation faible. On verra ainsi comment se répartissent les mot-outils en fonction de leur initiale, vocalique ou consonantique, et quels ils sont.

Dans un second temps, nous déterminerons quel est l'usage de ces mots-outils chez les prosateurs et chez les poètes dactyliques, de façon courante.

III. Vocabulaire constitutif de la catégorie des mots-outils dans les partages trochaïques marqués de ponctuation faible.

Nous limitons notre enquête aux cas de ponctuation faible, car c'est dans cet ordre de ponctuation que se manifeste le phénomène que nous nous proposons d'étudier avec la plus grande évidence. Rappelons qu'en cas de ponctuation forte, le nombre de mots-outils à initiale consonantique est assez peu différent du nombre des mots-outils à initiale vocalique; en cas de ponctuation faible, la proportion passe de 1 à 4.

Dans les 13920 vers analysés, les mots-outils se répartissent de la façon suivante, en fonction de la lettre initiale:

- initiale vocalique: *an, at, et, ibi, ita, ubi, ubique, ut, uti, utrum; ab, ad, in*. Au total, 11 coordonnants ou subordonnants et 3 prépositions.
- initiale consonantique: *nec, neque, nisi, quibus, quid, quis, quod, quot, sed, simul, tamen, uel; per*. Au total, 12 coordonnants et une préposition.

Nous avons déjà signalé que nous faisons rentrer parmi les mots-outils des adverbes dont l'usage équivaut fréquemment à celui d'un coordonnant. Les deux séries sont à peu près équivalentes en nombre: 14 et 13. Il n'est pas de raison, en principe, pour que les mots-outils à initiale vocalique soient employés plus fréquemment que les mots-outils à initiale consonantique; les *nec, neque, quod* ou *sed*, pourraient contrebalancer les *an, at, et* ou *ut*. D'ailleurs, dans une statistique plus générale, puisqu'elle ne tient pas compte de la quantité des syllabes, M.L. Nougaret a montré que les monosyllabes de liaison à initiale vocalique et ceux à initiale consonantique sont en nombre à peu près égal, tant dans les *Bucoliques* que dans le *De Signis* (NOUGARET 1966).

Pourtant les fréquences d'emploi des mots de ces deux séries sont très différentes dans le cas que nous étudions. Le tableau VIII donne, par pied, le nombre d'emplois relevés pour chaque mot dans le matériel étudié, soit 13920 hexamètres.

Il ressort de cette analyse verbale que c'est le très grand nombre des emplois de *et* qui fait que les liaisons CV sont plus nombreuses que les liaisons VC. La disproportion entre ces emplois de *et* en comparaison des autres mots-outils est telle que nous avons jugé inutile de donner la répartition de détail à travers les oeuvres, tant le phénomène est constant. Nous nous proposons d'ailleurs de mener une enquête particulière sur ces emplois de *et*, à partir d'analyses étendues. On notera de plus que, si l'on retranche les 194 *et* du total des mots-outils à initiales vocalique, soit 248-194, il reste 54, c'est à dire, à 4 unités près, le nombre total des mots-outils à initiale consonantique (58).

Etudiant les liaisons syllabiques entre II/III et III/IV, M.J. SOUBIRAN s'était trouvé devant le même problème; relevant une majorité de liaisons CV fondée sur l'emploi de monosyllabes de liaison à initiale vocalique, il concluait à une préférence affirmée par les poètes latins en faveur de ces mots, car "les dépouillements de M. Nougaret permettent d'avancer que, dans la prose de Cicéron comme dans les vers des *Bucoliques*, les monosyllabes de liaison sont à peu près aussi souvent à initiale vocalique qu'à initiale consonantique". (SOUBIRAN 1968:421)

Nous serions tenté de le suivre dans cette conclusion, sur le vu des résultats précédents. Mais en cas de partage trochaïque, les faits nous semblent beaucoup plus complexes et nous avons cru devoir reconsidérer dans son ensemble le problème de la répartition

des initiales vocaliques ou consonantiques dans la catégorie des mots-outils employés tant en prose que dans la poésie dactylique.

TABLEAU VIII. Nombre des emplois des mots-outils à initiale vocalique et des mots-outils à initiale consonantique.

Mots-outils	I	II	III	IV	V	Total
an	2	-	-	-	1	3
at	1	-	1	1	1	4
et	108	5	10	7	64	194
ibi	-	1	-	-	-	1
ita	-	1	1	-	-	2
ubi	3	-	1	-	3	7
ubique	1	-	-	-	-	1
ut	5	-	2	2	13	22
uti	-	3	-	-	1	4
utrum	1	1	-	-	1	3
ab	1	-	-	-	1	2
ad	1	-	-	-	2	3
in	1	-	-	-	1	2
	124	11	15	10	88	248
nec	4	-	2	-	2	8
neque	3	-	1	-	5	9
nisi	1	1	-	-	-	2
quibus	2	2	-	-	1	5
quid	-	-	-	-	2	2
quis	-	-	-	-	4	4
quod	3	1	-	1	4	9
quot	1	-	-	-	1	2
sed	4	1	1	-	5	11
simul	-	-	1	-	-	1
tamen	1	-	2	-	-	3
uel	-	-	-	-	1	1
per	-	-	-	-	1	1
	19	5	7	1	26	58

Note: Le total des mots-outils à initiale consonantique s'élève ici à 58, alors que pour la même catégorie, le relevé général donnait le chiffre de 60 (cf. tableau VII). C'est que nous n'avons pas compris dans le présent relevé deux cas tout à fait particuliers:

- *si* élidé sur initiale brève (Lucilius, 268);
- une liaison VC assurée par *-que* sur mot de sens plein employé en prolepse (Catulle, LXIV, 30). Cet exemple, unique en son genre, a été classé au relevé général dans la catégorie des mots-outils en tenant compte du fait que l'emploi du coordonnant *-que* pouvait constituer une contrainte lexicale dans la réalisation de la liaison syllabique. Les poètes évitent de faire débuter une proposition à un mode personnel après *-que*. Outre ce vers de Catulle, en 13920 hexamètres, nous n'avons trouvé que 4 autres exemples d'une telle succession, mais *-que* précède alors toujours un mot-outil (*Métamorphoses*, I, 739; Horace, *Satires* I, 1, 29; Perse, III, 117; *Argonautiques*, II, 166).

IV. Fréquence des mots-outils à initiale vocalique et des mots-outils à initiale consonantique assurant la coordination de propositions à un mode personnel ou la subordination, en poésie et en prose.

La première démarche consiste à déterminer les fréquences d'emploi des mots-outils à initiale vocalique et des mots-outils à initiale consonantique assurant coordination ou subordination de propositions à un mode personnel en prose d'une part, dans la poésie hexamétrique d'autre part, en quelque situation métrique que soient ces mots, et non pas dans les seuls cas de partage trochaïque. Mais deux conditions doivent être préalablement posées: premièrement, qu'il s'agisse d'un vocabulaire strictement comparable; deuxièmement, que les textes examinés soient de longueur égale.

Donc, afin d'éliminer toute ambiguité dans le classement, les relevés sont limités aux mots suivants:

- mots de liaison à initiale vocalique: *an, at, et, ubi, ut, uti, uter* (et ses diverses formes);
- mots de liaison à initiale consonantique: *nec, neque, nisi, quid, quis, quibus, quod, quot, sed, tamen.*

Il importait, en effet, de ne retenir que les mots usuels figurant dans la statistique établie ci-dessus et d'exclure les quelques mots qui, en prose, n'introduisent pas normalement une proposition à un mode personnel, puisque telle est la définition de la pause

de sens; soit les adverbes *ibi, ita, ubique, simul*, le coordonnant *uel* et les prépositions. Ces mots, pris en compte dans le tableau VIII sont très peu nombreux: 14 sur un total de 306 mots-outils.

Il est d'autre part nécessaire de se fonder sur des relevés comparables, donc d'analyser des textes de prose ayant à peu près la même étendue que les tranches de 1000 vers sur lesquelles nous avons travaillé. M.J. HELLEGOUARC'H a employé cette méthode et calculé que 6500 mots représentent la valeur d'un millier d'hexamètres; nous l'avons suivi sur ce point (HELLEGOUARC'H 1964:n.16 et n.2).

On a donc compté le nombre des emplois de chacun des mots des listes données ci-dessus dans le choix de textes suivant; pour confirmation du nombre moyen calculé par M.J. HELLEGOUARC'H, nous indiquons entre parenthèses le nombre réel de mots; dans chacune des tranches de mille vers, il y a, en moyenne, 6506 mots; si l'on y ajoute le relevé des *Bucoliques*, qui n'ont que 829 vers, en ramenant le nombre réel des mots comptés à un nombre fictif pour mille vers, cette moyenne passe à 6515 mots.

- Poésie: Lucrèce (6458), les *Bucoliques* (5475 mots pour 829 vers), les *Géorgiques* (6495), l'*Enéide* (6485), les *Satires* (6540) et les *Epîtres* (6587) d'Horace, les *Métamorphoses* d'Ovide (6609), la *Pharsale* (6308), la *Thébaïde* (6422), les *Argonautiques* (6667), les *Satires* de Juvénal (6494).
- Prose: 6500 mots environ de César, *Bellum ciuile* I; Cicéron, *Brutus, pro Milone, Correspondance* (éd. Constans, t.I, p.126-171); Salluste, *Catilina;* Tite-Live, III; Sénèque, *de Ira*, III, *ad Lucilium* 1-14; Pline le Jeune, *Panégyrique de Trajan;* Tacite, *Dialogue des orateurs, Annales*, XIII. Ces textes sont lus dans la Collection des Universités de France-Guillaume Budé, où la composition typographique est constante: 27 lignes pleines représentent en moyenne 200 mots, soit 878 lignes pour 6500 mots.

On a cru pouvoir se limiter à un choix parmi les oeuvres poétiques retenues pour l'étude générale du partage trochaïque étant donné qu'il s'agit d'apprécier un usage assez large, comprenant tous les emplois des mots-outils considérés, quelle que soit la quantité de la première syllabe. De plus, il eût été vain d'envisager ici les fragments d'Ennius et de Lucilius.

Les tableaux IX et X proposent les résultats de ces analyses. Ils donnent, dans l'ordre:

- le nombre de mots de liaison à initiale vocalique,
- le nombre de mots de liaison à initiale consonantique,
- le nombre total de mots de liaison employés.

TABLEAU IX. Mots de liaison à initiale vocalique et mots de liaison à initiale consonantique dans la poésie en hexamètres.

Auteurs/ Oeuvres	Initiale vocalique	Initiale consonantique	Total
Lucrèce	130	204	334
Buc.	119	92	211
Géorg.	158	83	241
Enéide	149	64	213
Hor. *Sat.*	173	111	284
Hor. *Epist.*	174	104	278
Métam.	137	97	234
Lucain	105	97	202
Théb.	101	94	195
Arg.	126	81	207
Juvénal	114	150	264
Total	1486	1177	2663

TABLEAU X. Mots de liaison à initiale vocalique et mots de liaison à initiale consonantique dans la prose

Auteurs/ Oeuvres	Initiale vocalique	Initiale consonantique	Total
Bell. ciu.	90	89	179
p.Milone	127	239	366
Brutus	139	139	278
CIC. *Corr.*	146	139	285
SALL. *Cat.*	59	130	189
Tite-Live	110	102	212
SEN. *de ira*	122	158	280
SEN. *ad Luc.*	143	152	295
PL. *Panég.*	138	191	329
TAC. *Or.*	117	151	268
TAC. *Ann.*	104	113	217
Total	1295	1603	2898

La comparaison des résultats donnés par les tableaux IX et X appelle les remarques suivantes:

Il y a dans les extraits d'oeuvres en prose un peu plus de mots de liaison que dans les extraits d'oeuvres poétiques: 2898 contre 2663, soit 235 en plus, environ 4 % du total. Mais si on regarde les faits en détail, par oeuvre, on s'aperçoit que les habitudes d'usage sont à peu de chose près comparables. Dans les oeuvres poétiques, le nombre des emplois varie entre 195 et 334; dans les oeuvres en prose, entre 179 et 336. Un classement en échelle fait ressortir la concordance entre chaque série d'oeuvres, les emplois poétiques s'incluant très exactement dans la liste des emplois en prose, et la différence entre les extrêmes étant de l'ordre de 16 à 32 unités.

Comparaison du nombre des emplois des mots de liaison entre prose et poésie.

pro Milone	366		
		334	Lucrèce
Panégyrique	329		
ad Lucilium	295		
CIC.*Corr.*	285	284	Horace, *Satires*
de Ira	280		
Brutus	278	278	Horace, *Epîtres*
Dial. Or.	268		
		264	Juvénal
		241	*Géorgiques*
		234	*Métamorphoses*
Ann. XIII	217	213	*Enéide*
Tite-Live	212	211	*Bucoliques*
		207	*Argonautiques*
		202	Lucain
		195	*Thébaïde*
Catilina	189		
Bell. ciu.	179		

Il semble que l'emploi des mots de liaison varie sensiblement suivant les genres littéraires: satires, épîtres et poésie didactique offrent le plus grand nombre d'emplois et par là se rapprochent des traités, des dialogues, des lettres et des oeuvres oratoires. Au contraire, l'épopée offre un usage plus réservé, mais constant: sur 200 mots de liaison environ, *Thébaïde, Pharsale, Argonautiques, Enéide,* ne diffèrent pas de plus de 8 emplois. En regard, en prose, se classent les oeuvres historiques.

L'homogénéité de ces résultats, tant dans le dénombrement que dans le groupement des genres littéraires, incite à penser que nous ne sommes pas en présence de faits de hasard et que ces analyses faites sur des fragments d'oeuvre rendent compte d'une tendance générale de la langue, valable pour le mode d'expression le plus courant, la prose, comme pour le mode le plus recherché, la poésie.

Il n'en est plus de même dans la répartition entre mots de liaison à initiale vocalique et mots de liaison à initiale consonantique.

Les initiales vocaliques sont plus fréquentes en poésie: 191 de plus qu'en prose (1486 - 1295 = 191).

En revanche, les initiales consonantiques sont beaucoup moins fréquentes qu'en prose (426 de moins en poésie (1603 - 1177 = 426).

De plus, si l'on considère, dans chaque série d'oeuvres, la répartition des initiales, on s'aperçoit:

- qu'en poésie, les mots de liaison à initiale vocalique sont plus nombreux que les mots de liaison à initiale consonantique de 309 unités (1486 - 1177 = 309).

- qu'en prose, la situation est inverse; les mots de liaison à initiale consonantique sont plus nombreux que les mots de liaison à initiale vocalique de 308 unités (1603 - 1295 = 308).

Si, enfin, on regarde comment se présentent les fréquences d'emploi à travers les oeuvres particulières, on constate que ces faits sont pratiquement constants, donc attachés au mode d'expression.

Ainsi, dans 9 oeuvres poétiques sur 11, les mots de liaison à initiale vocalique sont plus nombreux que les mots de liaison à initiale consonantique (de 7 à 85 unités en plus); seules les oeuvres extrêmes chronologiquement présentent une situation inverse (Lucrèce, 74 initiales consonantiques de plus que d'initiales vocaliques; Juvénal, 36).

Au contraire, dans 7 oeuvres de prose sur 11, les mots de liaison à initiale consonantique sont plus nombreux que les mots de liaison à initiale vocalique (de 9 à 112 unités en plus); dans les

4 autres oeuvres, les emplois sont équivalents, ou à peu près: 8 initiales vocaliques de plus que d'initiales consonantiques chez Tite-Live, 7 dans les *Lettres* de Cicéron, une seule en plus dans le *Bellum ciuile,* égalité parfaite dans le *Brutus.*

Résumons les constatations faites:

- l'usage général des mots de liaison est en nombre à peu près le même en prose et en poésie;
- les mots de liaison à initiale vocalique sont employés plus fréquemment dans les oeuvres poétiques quand il s'agit d'introduire une proposition à un mode personnel, c'est à dire quand une pause de sens intervient devant eux;
- les mots de liaison à initiale consonantique sont, employés dans les mêmes conditions, plus nombreux dans les oeuvres de prose, qui tendent, en cas limite, à présenter, au plus, autant de mots de liaison à initiale vocalique que de mots de liaison à initiale consonantique.

Rappelons qu'il a été tenu compte des mots de liaison à quelque place qu'ils se situent dans le vers et quelle que soit la quantité de la syllabe initiale; ainsi, ont été pris en compte les mots de liaison employés au début du vers, bien que ne se pose pas alors le problème de la liaison syllabique. Nous avons cependant établi la statistique sur cette base afin de ne pas limiter la comparaison avec l'usage de la prose.

Mais si l'on serre l'analyse de plus près et qu'on prenne seulement en considération les mots de liaison situés à l'intérieur du vers, donc ceux sur lesquels se réalisent les liaisons syllabiques, il est significatif de constater que les initiales vocaliques, non seulement sont toujours plus nombreuses que les initiales consonantiques, mais encore que la différence entre ces deux catégories reste exactement la même que dans l'analyse précédente, où étaient comptés les mots placés en tête du vers, ce que montre le tableau XI.

TABLEAU XI. Mots de liaison à initiale vocalique et mots de liaison à initiale consonantique à l'intérieur de l'hexamètre.

Auteurs/ Oeuvres	Initiale vocalique	Initiale consonantique	Total
Lucrèce	94	134	228
Buc.	72	60	132
Géorg.	116	33	149
Enéide	122	42	164
Hor. *Sat.*	153	87	240
Hor. *Epist.*	145	69	214
Métam.	100	66	166
Lucain	76	77	153
Thébaide	90	79	169
Argonaut.	108	62	170
Juvénal	82	140	222
Total	1158	849	2007

Il n'y a plus qu'à comparer les résultats obtenus dans les deux cas:
- Sont pris en considération tous les mots de liaison, quelle que soit leur place dans le vers (cf. tableau IV): il y a 309 initiales vocaliques de plus que d'initiales consonantiques:
1486 - 1177 = 309.
- sont pris en considération les seuls mots de liaison situés à l'intérieur du vers (cf. tableau IV): il y a toujours 309 initiales vocaliques de plus que d'initiales consonantiques:
1158 - 849 = 309.

La différence reste la même entre des nombres dissemblables: dans chacune des catégories considérées (initiales vocaliques et initiales consonantiques) a été retranché le même nombre de cas, soit 328. Faisons en effet la différence entre le nombre de mots de liaison à initiale vocalique relevé dans notre premier calcul (totalité des emplois, tableau IX) et le nombre relevé dans le second calcul (mots situés à l'intérieur du vers seulement, tableau XI), nous obtenons: 1486 - 1158 = 328. Opérons de même pour les mots de liaison à initiale consonantique, le résultat est identique: 1177 - 849 = 328.

Ce qui revient à dire qu'en tête de vers, il y a exactement le même nombre de mots de liaison à initiale vocalique que de mot

liaison à initiale consonantique. Nous dirons donc que si, dans les textes poétiques fragmentaires examinés, les mots de liaison à initiale vocalique sont plus nombreux que les mots de liaison à initiale consonantique, cela est dû aux emplois de ces mots à l'intérieur du vers, là-même où se réalisent les liaisons syllabiques, et qu'il peut y avoir corrélation entre la fréquence des initiales vocaliques et le fait de liaison syllabique.

Mais cette façon de prendre le problème réunit, sans distinction, tous les cas de liaison syllabique, quel que soit l'intermot concerné, quand intervient une ponctuation marquant une pause de sens faible. Aussi, émettrons-nous une seconde hypothèse: le phénomène constaté à propos des partages trochaïques fait partie d'un ensemble plus vaste dans lequel il serait une manifestation exceptionnelle, tout se passant comme si la situation des mots de liaison à la seconde brève des partages trochaïques marqués d'une pause de sens faible rendait encore plus nécessaire la liaison CV. En effet, à en juger, à titre indicatif, sur les pourcentages généraux établis, dans le cas le plus général, les initiales vocaliques représentent de 56 à 58 % des mots de liaison employés (prise en compte de tous les mots de liaison, 1486 initiales vocaliques sur 2663 mots de liaison = 56 %; cf. tableau IX. Prise en compte des mots de liaison situés à l'intérieur du vers, 1158 initiales vocaliques sur 2007 mots de liaison = 58 %; cf. tableau XI). En cas de partage trochaïque, ces mêmes initiales vocaliques représentent 81 % des mots de liaison employés et à eux seuls, les emplois de *et* comptent pour 63 % du total (248 initiales vocaliques sur 306 mots de liaison = 81 % - 194 *et* sur 306 mots de liaison = 63 %; cf. tableau VIII).

Ainsi, ce ne serait pas un hasard si, pour coordonner des propositions à un mode personnel, les poètes emploient plus de mots de liaison à initiale vocalique que les prosateurs et moins de mots de liaison à initiale consonantique , alors que poètes et prosateurs emploient à peu près le même nombre de mots de liaison dans des textes d'étendue équivalente; ce ne serait pas un hasard si cet usage s'affirme uniquement à l'intérieur de l'hexamètre. Faut-il reconnaître ici les nécessités de l'*oratio uincta*, se manifestant parfois sur un plan très général où les problèmes de quantité des syllabes deviennent secondaires; nécessités plus rigoureuses quand joue la quantité brève, dans les cas de partage trochaïque marqué de ponctuation?

Conclusion

A ce stade de notre étude et pour conclure, nous avancerons avec prudence un certain nombre d'explications: on voudra bien les prendre pour des suggestions et non des conclusions arrêtées, puisqu'en grande partie nous avons fondé nos analyses sur des textes fragmentaires. Mais, nous l'avons dit, en entreprenant cette enquête qui prend en considération un nombre important d'oeuvres poétiques, notre intention était surtout d'éprouver des méthodes, de définir des problèmes et d'ouvrir des perspectives.

La liaison syllabique CV est bien recherchée en poésie quand le mot qui fournit la seconde brève du partage trochaïque marqué de ponctuation est un outil de liaison, ce qui n'est pas sans conséquence sur le choix de ces mots et l'usage qui en est fait à l'intérieur même du vers.

Mais aussi, considérant l'ensemble des faits, n'étions-nous pas tout à fait dupe d'une apparence en découvrant qu'en cas d'intervention d'une ponctuation forte ou faible entre les brèves du dactyle, les poètes réalisaient plus volontiers la liaison CV, en valeur absolue (58 % de liaisons CV contre 42 % de liaisons VC; cf. tableau IV), et en valeur relative, si l'on compare avec les liaisons réalisées quand aucune pause n'intervient à l'intermot considéré (26 % de liaisons CV contre 74 % de liaisons VC; cf. tableau III). Cela signale que, dans l'enchaînement ordinaire d'un discours lié par d'exceptionnelles conventions métriques et phonétiques, le fait était ressenti comme dominant et nécessaire. Prenons-en pour indice complémentaire ce qu'on a pu observer à propos des seuls partages trochaïques marqués de ponctuation forte, sur lesquels nous avons d'ailleurs pu faire des analyses exhaustives. En valeur absolue, les liaisons VC sont plus nombreuses que les liaisons CV qui ne représentent plus que 45 % des partages (cas de ponctuation trochaïque forte dans la totalité des oeuvres étudiées; cf. tableau VI). Cependant, ces 45 % sont un pourcentage encore bien plus élevé que les 26 % qui représentent, dans les partages trochaïques non marqués de ponctuation, la fréquence d'usage de cette même liaison CV.

Il reste à expliquer la divergence des résultats obtenus lorsqu'on distingue dans l'analyse, d'une part les mots de sens plein, d'autre part les mots-outils. Nous venons de voir que l'usage des

mots-outils à initiale vocalique permet de réaliser, à l'intérieur du vers, et plus encore en cas de partage trochaïque marqué de ponctuation, des liaisons CV en plus grand nombre que les liaisons VC réalisées sur mots-outils à initiale consonantique. Pourquoi, en revanche, la liaison VC est-elle plus fréquente que la liaison CV si l'on ne considère que les mots de sens plein? Rappelons en effet les pourcentages caractéristiques: pour cette catégorie de mots, 70 % de liaisons VC dans les partages trochaïques marqués de ponctuation forte ou faible (cf. tableau V); 68 % dans les partages marqués de ponctuation forte (analyse sur la totalité des oeuvres; cf. tableau VI); 75 % dans les partages trochaïques marqués de ponctuation faible (cf. tableau VII).

Il faut peut-être chercher l'explication dans la figure métrique elle-même et penser à la contradiction essentielle qu'elle recèle: une séparation de mots provoquant la séparation des brèves du dactyle, créant une gêne dans la diction poétique, plus ou moins accusée selon que la pause exigée par le sens prend plus ou moins d'importance; difficulté qui ne se résoudrait que par un artifice de diction restaurant l'isochronie d'un demi-pied comprenant deux brèves et "quelque chose en plus", qui devrait être une suspension du discours, si légère soit-elle; d'où une accélération possible du débit. Les mots de liaison s'accordaient bien avec cette exigence; marquant à la fois rupture du sens et rupture du rythme, ils annonçaient nettement la séparation entre des membres syntaxiques différents; on pouvait alors donner préférence à ceux d'entre eux qui, parce qu'ils avaient une initiale vocalique, rendaient plus aisé un enchaînement rapide en réalisant une liaison CV, en accord avec l'accélération du débit que semble postuler le partage des brèves.

Il n'en allait plus de même avec les mots de sens plein, qui ne présentent pas les mêmes avantages que les mots de liaison pour marquer l'orientation nouvelle du sens, le passage d'une phrase à l'autre. Quand l'un d'eux fournit la seconde brève d'un partage trochaïque marqué de ponctuation, ouvrant ainsi un membre syntaxique, il succède normalement à un mot de même catégorie, qui fournit la brève du partage et clôt le membre syntaxique précédent. Ce vocabulaire n'est pas apte à signaler une telle distribution syntaxique. Or comme le partage trochaïque suppose, par nature, une accélération du débit, qu'adviendrait-il si se réalisait alors une liaison CV facilitant l'enchaînement d'un mot sur l'autre et, pour ainsi dire, les confondant phonétiquement? On obtiendrait une continuité du discours, effaçant si bien la couture réalisée entre les mots par le rythme poétique qu'il deviendrait difficile à l'oreille de saisir, à quelque signe que ce fût, l'organisation du sens. La liaison VC, au contraire, soulignant la séparation entre les mots, convient à une telle situation et remédie à son tour à l'enchaînement assuré par la diction particulière qu'appelle le partage des brèves. Les deux aspects du phénomène, en apparence contradictoire, sont complémentaires.

Il n'y a rien qui s'oppose aux analyses faites par M.J. Soubiran et aux explications qu'il en a tirées, car celles-ci concernent les intermots situés entre des pieds successifs. On imagine facilement qu'une pause de sens intervenant entre deux pieds risque de perturber le cours normal de l'hexamètre, de faire croire à une division artificielle du vers, que la diction rythmée par l'ictus ne ferait qu'accentuer: d'où, en ces cas, la préférence accordée à la liaison CV, à quelque catégorie de vocabulaire qu'appartiennent les mots qui suivent cette pause.

La situation du partage trochaïque est toute différente; la division des brèves du dactyle est déjà un phénomène phonétique intéressant beaucoup plus la structure du pied que la structure du vers. Les problèmes d'articulation ne sont pas les mêmes qu'ailleurs, parce qu'ils se posent en un point précis où tout accident risque de contrevenir à la régularité d'un rythme par nature fugace. Ainsi, l'intervention d'une pause de sens dans le partage des brèves apparaît comme un fait métrique original, disons comme une "figure" où se réalise une synthèse des moyens d'expression. L'art du poète consiste précisément à tirer parti des difficultés techniques ainsi créées, de telle sorte qu'en les surmontant, par une subtile alliance du rythme et de la langue, il achève la création poétique.

BIBLIOGRAPHIE

GERARD, J., La ponctuation trochaïque dans l'hexamètre latin d'Ennius à Juvénal. Recherches sur les modalités de l'accord entre la phrase et le vers, thèse d'Etat. Paris, Les Belles Lettres 1980

HELLEGOUARC'H, J., Le monosyllabe dans l'hexamètre latin. Essai de métrique verbale. Paris, Klincksieck 1964

NOUGARET, L., Analyse verbale comparée du De Signis et des Bucoliques. Paris, Les Belles Lettres 1966

PERRET, J., Mots et fins de mots trochaïques dans l'hexamètre latin. Dans: Revue des Etudes latines, 32, 1954, 183-199

PERRET, J., Le partage du demi-pied dans les anapestiques et dans l'hexamètre. Dans: Revue des Etudes latines, 33, 1955, 352-366

SOUBIRAN, J., L'élision dans la poésie latine. Paris, Klincksieck 1966 (= 1966a)

SOUBIRAN, J., Ponctuation bucolique et liaison syllabique en grec et en latin. Dans: Annales de la Faculté des Lettres de Toulouse, Pallas, 8, 1966, 21-52 (= 1966b)

SOUBIRAN, J., L'hexamètre latin. Problèmes de structure et de diction. Dans: Revue des Etudes latines, 46, 1968, 410-424

SOUBIRAN, J., Pauses de sens et cohésion métrique entre les pieds médians de l'hexamètre latin. Dans: Annales de la Faculté des Lettres de Toulouse, Pallas, 16, 1969, 107-151

A Note on Jean Gérard's Paper "Liaisons syllabiques dans les partages trochaïques de l'hexamètre latin"

Rüdiger Grotjahn, Bochum

0. Introduction

In his article "Liaisons syllabiques dans les partages trochaïques de l'hexamètre latin" (pp. 168-211 in this volume), J. GÉRARD presents a great deal of very interesting data from numerous Latin authors. These data are analyzed with the help of certain descriptive statistics and a number of interesting hypotheses are presented. In the following short note I will show how these hypotheses, as well as some additional hypotheses, can be tested by applying relatively simple methods of inferential statistics. In this way I hope to demonstrate that the use of inferential statistics leads to more reliable conclusions about the problem of syllabic liaison in trochaic division in Latin hexameter verse. Readers interested in more sophisticated techniques of analyzing data such as those presented by GÉRARD are referred to e.g. BISHOP/FIENBERG/HOLLAND (1975) or REYNOLDS (1977).

1. Analysis of GÉRARD's Table 1

On page 10, referring to his Table 1, GÉRARD states that there are apparent differences among different authors in the use of weak trochaic punctuations, as well as in the frequencies of weak trochaic punctuations, in the various different verse feet. There seems to be a tendency towards an increased use of weak trochaic punctuations, especially in feet I and V.

In order to test this latter hypothesis, we first calculate the frequency which we would expect under the hypothesis of equidistribution for each foot. In Catullus, for instance, GERARD has found 8 weak trochaic punctuations. Thus we would ex-

pect 8/5 = 1.6 weak trochaic punctuations in each foot. Since the sample from Catullus consists of 408 verses, the number of verse feet analyzed by GÉRARD is 5(408) = 2040. Since compared to n = 2040 the frequencies of weak trochaic punctuations observed in Catullus are very small, we can use the Poisson distribution to test whether the observed frequencies differ significantly from the expected ones. In the case of foot III in the sample from Catullus, for instance, the observed frequency is 5. Using tables of the Poisson distribution we find that

$$P(X \geq 5;\ \lambda = 1.6) = \sum_{x=5}^{\infty} \frac{e^{-1.6}1.6^x}{x!} = 0.023682.$$

In the same way we obtain e.g. for the Aeneid $\hat{\lambda} = 14/5 = 2.8$. In foot IV, for instance, the observed frequency is 1 and thus smaller than $\hat{\lambda}$. We therefore compute

$$P(X \leq 1;\ \lambda = 2.8) = \sum_{x=0}^{1} \frac{e^{-2.8}2.8^x}{x!} = 0.231078.$$

We now adopt the following convention: If the observed frequency is the same as the expected one, we use the sign 0. In case the observed frequency is smaller than the expected one and if $P > 0.025$, we use the sign (-); if $0.005 < P \leq 0.025$, the sign -; if $0.0005 < P \leq 0.005$, the sign --; and if $P \leq 0.0005$, the sign ---. If the observed frequency is greater than the expected one, by analogy we use (+), +, ++, and +++. In this way, with the help of the Poisson distribution, we get a pattern of signs for each verse foot and for each author/work in GÉRARD's Table 1 which makes the regularities and irregularities in the use of weak trochaic punctuations in the different feet and/or among the different authors/works easily visable. It should, however, be noted that for several reasons the procedure adopted is only approximative.

The result of the analysis is presented in Table 1.

Table 1. Analysis of the Deviations of the Observed Frequencies from the Expected Frequencies for the Data of GÉRARD's Table 1

Authors/ Works	Weak Trochaic Punctuations					
	I	II	III	IV	V	$\hat{\lambda}$
Ennius	O	O	O	O	O	1
Lucilius	(-)	(+)	(+)	(-)	(-)	2.2
Catullus	(-)	(+)	+	(-)	(-)	1.6
Lucretius	(+)	(-)	(-)	--	++	8.2
Buc.	(+)	-	(+)	(-)	(+)	3.6
Georg.	(+)	(-)	(-)	--	++	6
Aeneid	+	(-)	(+)	(-)	(-)	2.8
Hor. *Sat.*	(-)	--	(-)	--	+++	11.2
Hor. *Epist.*	O	(-)	(-)	(-)	+++	10
Metam.	+++	-	--	--	(+)	5.6
Lucan	+++	(-)	(-)	-	(+)	4.6
Persius	(+)	(-)	-	-	+++	6.4
Silvae	+++	-	-	-	(+)	5
Thebais	+++	--	--	--	+	6
Arg.	+++	-	--	-	(-)	6
Juvenal	+	--	(-)	--	++	5.4
Total	+++	---	---	---	+++	85.6

Table 1 clearly shows that, especially from Horace's *Satires* to Juvenal, weak trochaic punctuations are far more frequently used in the first and/or fifth feet than in the remaining ones.

We now want to test whether the variables 'kind of author/ work' and 'frequency of weak trochaic punctuations in different verse feet' correlate. In order to take into account that the frequencies in GÉRARD's Table 1 are calculated from samples consisting of 5(408) = 2040 to 5(1000) = 5000 verse feet, we arrange the data as shown in Table 2.

Carrying out a χ^2 test we obtain X^2 = 243.53 with df = = (6-1)(16-1) = 75, a highly significant value. But if we compute Cramér's contingency coefficient C_c (cf. GROTJAHN in this volume p. 101), we get C_c = 0.0265, which seems to indicate that the degree of association in GÉRARD's Table 1 is very weak. In

Table 2. Contingency Table Constructed from GÉRARD's Table 1

Authors/ Works	I	II	III	IV	V	Other	Total
Ennius	1	1	1	1	1	2135	2140
Lucilius	1	4	5	-	1	3014	3025
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
Juvenal	11	-	3	-	13	4963	5000
Total	165	30	49	14	170	69172	69600

interpreting this result one should, however, take into consideration that C_c is computed from an extremely large sample and that furthermore C_c reflects the *average* association in GÉRARD's Table 1. Hence, there may be a far stronger association between particular rows and columns .

We now test each cell of GÉRARD's Table 1 as to whether its contribution to the X^2 value obtained can be regarded as significant. (For an explanation of this test, which is known as 'Test for deviation in a single cell' or 'Test of the (standardized) residuals', see HABERMAN 1973, ŘEHÁK/ŘEHÁKOVÁ 1980 and GROTJAHN in this volume pp. 97ff). We use the normal approximation (formulas 2a and 2b on p. 99 in this volume) if the expected frequency is greater than 10; otherwise we use the Poisson approximation.

For the fifth foot in the sample from Horace's *Satires*, for example, we get the expected frequency $E(n_{85})$ = 170(5000)/69600 = = 12.2. We thus apply the normal approximation and obtain z = = 5.5870. Since $P(Z \geq 5.5870) < 0.0005$, we mark it +++. If we compute the expected frequency for the fifth foot in the sample from Lucilius, we get $E(n_{25})$ = 170(3025)/69600 = 7.4. We thus use the Poisson distribution and calculate $P(X \leq 1; \lambda = 7.4)$ = = 0.005135. Since $0.005 < P \leq 0.025$, we mark it -.

The results of the test for deviation in a single cell are shown in Table 3.

As can be seen immediately from Table 3 it is primarily verse foot V which contributes to the association between the two vari-

Table 3. Test for Deviation in a Single Cell for the Data from GÉRARD's Table 1

Authors/ Works	Weak Trochaic Punctuations					
	I	II	III	IV	V	Total
Ennius	(-)	(+)	(-)	(+)	(-)	-
Lucilius	-	(+)	+	(-)	-	(-)
Catullus	(-)	(+)	+	(-)	-	(-)
Lucretius	(-)	++	(+)	(-)	(+)	(+)
Buc.	(-)	(-)	(+)	(+)	-	(-)
Georg.	(-)	(-)	(+)	(-)	(+)	(-)
Aeneid	(-)	(-)	(-)	(-)	--	--
Hor. *Sat.*	(-)	(+)	++	(+)	+++	+++
Hor. *Epist.*	(-)	(+)	(+)	+++	+++	+++
Metam.	(+)	(-)	(-)	(-)	(-)	(-)
Lucan	(+)	(-)	(-)	(-)	(-)	(-)
Persius	(+)	(+)	(-)	(+)	+++	++
Silvae	(+)	(-)	(-)	(-)	(-)	(-)
Thebais	(+)	(-)	(-)	(-)	(-)	(-)
Arg.	+++	(-)	(-)	(-)	-	(-)
Juvenal	(-)	(-)	(-)	(-)	(+)	(-)

ables under investigation. There are furthermore considerable differences between the various authors/works. Whereas e.g. in the *Aeneid* all observed frequencies are smaller than expected (though only one deviation is significant), in Horace the observed frequencies are significantly higher than expected in four of the five verse feet. Since in Horace the distribution of weak trochaic punctuations seems to be highly exceptional, we carry out a χ^2 test on Horace and the pooled remaining authors/works. This time we take only the number of weak trochaic punctuations into account and not the total number of verse feet analyzed by GÉRARD. The corresponding data are presented in Table 4.

Table 4. Frequencies of Weak Trochaic Punctuations in Horace and 14 Other Authors/Works

	I	II	III	IV	V	
Horace	20	7	15	8	56	106
Other	145	23	34	6	114	322
	165	30	49	14	170	428

The χ^2 test now yields $X^2 = 29.06$, which is highly significant ($P(\chi^2_4 \geq 29.06) < 0.0005$). If we compute C_c, we get $C_c =$ $= 0.2606$, which indicates a moderate degree of association and thus also a moderate difference between Horace and the pooled remaining authors/works.

If we want to test whether there is, for instance, a difference between the frequencies in the first foot in *Argonautica* and in Horace's *Satires*, we can proceed as follows: we first assemble the data as shown in Table 5.

Table 5. Comparison of *Argonautica* and Horace's *Satires*

	I	II-V	Total
Hor. *Sat.*	23	7	30
Arg.	10	46	56
	33	53	86

We then carry out a χ^2 test and obtain $X^2 = 28.57$ which corresponds to a probability of far less than 0.0005. Consequently there is a highly significant difference between the two frequencies. If we compute the φ coefficient (for a discussion of this coefficient see GROTJAHN, pp. 62ff in this volume), we obtain $\varphi = 0.5764$, which indicates a relatively strong difference between the two authors/works.

Finally I would like to suggest and test a hypothesis not explicitly stated by GÉRARD. If we look at the frequencies of weak trochaic punctuations in the first verse foot, the frequencies seem to increase according to the chronological order of the corresponding authors/works. In order to test this hypothesis, we first assign ranks to the authors/works: rank 1 to Ennius, rank 2 to Lucilius and so forth until Juvenal (rank 16). We then convert the absolute frequencies into relative ones by dividing each frequency by the corresponding number of verses. We then assign ranks to the relative frequencies as well. The data obtained are shown in Table 6.

Table 6. Verse Foot I: Chronological Rank Order of Authors/ Works vs. Rank Order of Relative Frequencies (%)

Authors/ Works	Rank	Frequencies %	Rank
Ennius	1	0.23	2
Lucilius	2	0.17	1
Catullus	3	0.25	3
Lucretius	4	0.90	6.5
Buc.	5	0.72	4
Georg.	6	0.90	6.5
Aeneid	7	0.80	5
Hor. *Sat.*	8	1.00	8.5
Hor. *Epist.*	9	1.00	8.5
Metam.	10	1.60	13
Lucan	11	1.40	11
Persius	12	1.54	12
Silvae	13	1.80	14.5
Theb.	14	1.80	14.5
Arg.	15	2.30	16
Juvenal	16	1.10	10

We now compare the two rank orders in Table 6 with the help of Spearman's rank correlation coefficient and obtain $r_s = 0.9079$. Since this value is highly significant ($P << 0.0005$), the hypothesis that the frequency of weak trochaic punctuations in the first foot tends to increase chronologically from Ennius to Juvenal is confirmed. Furthermore, as $r_s = 0.9079$ indicates, this tendency is very strong.

2. Analysis of GÉRARD's Table 2

As with GÉRARD's Table 1, I first test the deviation from equidistribution of the frequencies in the 5 different feet with the help of the Poisson model. The test results are shown in Table 7.

We see from Table 7 that although only a small percentage of the deviations are significant, there is a clear tendency toward the use of strong trochaic punctuations in foot III and in particular in foot V. If we compare the different authors/works, the patterns of *Bucolica*, Horace's *Satires* and Juvenal are especially striking. In particular Horace's *Satires* show very marked deviations from the hypothesis of equidistribution, as was the case with weak trochaic punctuations (cf. Table 1). If we compare the *direction* of the deviations in the corresponding cells of Table 7 and Table 1, we find a high number of coincidences of direction in the fourth and fifth feet (93% and 92%), and a relatively low number of coincidences, especially in verse foot I (50%).

Table 7. Analysis of the Deviations of the Observed Frequencies from the Expected Frequencies for the Data of GÉRARD's Table 1

Authors/ Works	Strong Trochaic Punctuations I	II	III	IV	V	$\hat{\lambda}$
Ennius	(-)	(+)	(+)	(-)	(-)	0.4
Lucilius	(-)	(-)	(+)	(+)	0	1.0
Catullus	0	0	0	0	0	0
Lucretius	(-)	(-)	(+)	(-)	(+)	0.6
Buc.	(-)	(-)	+++	--	(+)	5.4
Georg.	0	0	(+)	-	(+)	4.0
Aen.	(-)	(+)	(+)	(-)	(+)	1.6
Hor. *Sat.*	---	-	+	--	+++	8.4
Hor. *Epist.*	(-)	(-)	(+)	(-)	(+)	4.2
Metam.	(+)	(-)	(-)	(-)	(+)	2.0
Lucan	(-)	(-)	(-)	(-)	+	0.8
Persius	(-)	(+)	(-)	(-)	+	2.0
Silvae	++	(-)	(-)	(-)	0	2.0
Theb.	(+)	(-)	(-)	(-)	+	2.0
Arg.	(+)	(-)	(+)	(-)	(-)	1.4
Juvenal	-	-	-	-	+++	2.4
Total	(-)	--	++	---	+++	38.2

We now want to test whether the observed frequencies in GÉRARD's Table 2 deviate from the expected frequencies estimated in the same way as in Table 3. Since this time all expected frequencies are very small, a χ^2 test does not seem appropriate. Nevertheless, in order to be able to compare the use of strong and weak trochaic punctuations, we analyze the deviations in the same way as in Table 3. This time, however, because of the small expected frequencies, we use the Poisson distribution in all cases. The result of the analysis is presented in Table 8.

Table 8. Test for Deviation in a Single Cell for the Data from GÉRARD's Table 2

Authors/ Works	Strong Trochaic Punctuations					
	I	II	III	IV	V	Total
Ennius	(-)	(+)	(-)	(-)	(-)	(-)
Lucilius	(-)	(-)	(-)	+	(-)	(-)
Catullus	(-)	(-)	(-)	(-)	(-)	--
Lucretius	(-)	(-)	(-)	(-)	-	--
Buc.	(-)	(+)	+++	(-)	(+)	+++
Georg.	(+)	(+)	(+)	(-)	(+)	(+)
Aen.	(-)	(+)	(-)	(-)	(-)	(-)
Hor. *Sat.*	(-)	(+)	+++	(+)	+++	+++
Hor. *Epist.*	(+)	(+)	(+)	(+)	(+)	(+)
Metam.	(+)	(-)	(-)	(-)	(-)	(-)
Lucan	(-)	(-)	-	(-)	(-)	--
Persius	(-)	(+)	(-)	(-)	(+)	(+)
Silvae	++	(-)	-	(-)	(-)	(-)
Theb.	(+)	(-)	(-)	(-)	(+)	(-)
Arg.	(+)	(-)	(-)	(-)	-	(-)
Juvenal	(-)	(-)	-	(-)	(+)	(-)

We see from Table 8 that there are three authors/works (Catullus, Lucretius and *Metamorphoses*) with highly significant negative deviations and only two authors/works (*Bucolica* and Horace's *Satires*) with very highly significant positive deviations. Since the distributions of strong trochaic punctuations in *Bucolica* and Horace's *Satires* seem to be quite exceptional compared to the remaining authors/works, we carry out a χ^2 test on the pooled data from *Bucolica* and Horace's *Satires* and the pooled data from the remaining authors/works. If we take n = 191 as total sample size, we obtain $X_4^2 = 25.80$ and $C_c = 0.3675$. Hence the difference between *Bucolica* and Horace's *Satires* on the one hand and the remaining authors/works on the other hand is highly significant, although, as $C_c = 0.3675$ indicates, the difference is not very large.

3. *Analysis of GÉRARD's Tables 3 and 4*

We first want to test whether there is a difference between hexameter verse and prose in the use of liasons CV and VC in trochaic divisions not marked by punctuation. To this end we compare each author/work in GÉRARD's Table 3 to Cicero's *Second Philippica* with the help of a χ^2 test, where GÉRARD has found 1517 liaisons CV and 2575 liaisons VC. For Ennius, for instance, we get the following contingency table:

Table 9. Comparison of the Frequencies of Liaisons CV and Liaisons VC in Ennius and Cicero's *Second Philippica*

Authors/ Works	CV	VC	
Ennius	64	268	332
Phil. II	1517	2575	4092
	1581	2843	4424

The χ^2 test yields $X^2 = 42.34$. Hence there is a highly significant difference between the corresponding frequencies in Ennius and *Second Philippica*.

The results of all 16 χ^2 tests together with the corresponding φ coefficients are presented in Table 10.

As can be seen immediately from the X^2 values in Table 10, all 16 authors/works differ significantly from the *Second Philippica*. The differences are, however, in all cases very small, as is shown by the φ coefficients. In some cases, as for instance in that of Horace's *Satires* and *Epistles*, there seems to be almost no difference between verse and prose. As the sign of φ indicates, in all cases the relative frequencies of liaisons CV are smaller and the relative frequencies of liaisons VC greater than the corresponding values in the *Second Philippica*.

We now want to compare the observed frequencies in GÉRARD's Tables 3 and 4 with the frequencies expected under the hypothesis of independence. Since we can assume that because of the

Table 10. Comparison of the Frequencies of Liaison CV and Liaison VC in Hexameter Verse and Prose

Authors/Works	χ^2	φ
Ennius	42.34	-0.10
Lucilius	8.41	-0.04
Catullus	41.00	-0.10
Lucretius	61.75	-0.11
Buc.	11.76	-0.05
Georg.	56.16	-0.11
Aen.	92.24	-0.14
Hor. *Sat.*	6.32	-0.04
Hor. *Epist.*	8.14	-0.04
Metam.	52.53	-0.10
Lucan	135.43	-0.16
Persius	25.20	-0.07
Silvae	86.24	-0.13
Theb.	93.58	-0.13
Arg.	114.04	-0.15
Juvenal	15.28	-0.06
Total	258.97	-0.13

very large n at least some of the observed frequencies will deviate significantly from their expected values, we will not carry out a χ^2 test but immediately test the deviations in the same way as we did in constructing Tables 3 and 8. Neglecting the elisions in GÉRARD's Table 4 (because of their very low frequencies) we get the test results displayed in Table 11.

As can be seen from Table 11, due to the relatively large frequencies involved, many of the liaisons CV and VC not marked by punctuation show highly significant deviations from their expected values. In the case of syllabic liaisons marked by punctuation, however, there are only some significant deviations. Again, the data from Horace, which show highly significant positive deviations in the frequencies of syllabic liaisons marked by punctuation, are quite exceptional.

We now want to test the following hypothesis not explicitly stated by GÉRARD: "The higher the frequency of liaisons CV in a specific author/work, the higher the frequency of liaisons VC in that author/work". To test this hypothesis we first transform the absolute frequencies into relative ones, then we assign ranks

Table 11. Test for Deviation in a Single Cell for the Data from GERARD's Tables 3 and 4

Authors/Works	No Punctuation				Strong Punctuation			Weak Punctuation		
	El.	CV	VC	Total	CV	VC	Total	CV	VC	Total
Ennius	(-)	--	(-)	--	(-)	(-)	(-)	(-)	(-)	-
Lucilius	+++	--	---	---	(+)	(-)	(-)	-	(-)	(-)
Catullus	(-)	---	---	---	(-)	(-)	--	---	(+)	(-)
Lucretius	+++	---	---	---	-	-	--	(+)	+	(+)
Buc.	+++	+++	-	(+)	+	++	+++	(-)	(-)	(-)
Georg.	(-)	(-)	(-)	--	(+)	(+)	(+)	(+)	(-)	(-)
Aeneid	(-)	(-)	+++	+	-	(-)	(-)	-	-	--
Hor. *Sat.*	(+)	(+)	---	---	+++	+++	+++	+++	++	+++
Hor. *Epist.*	--	+++	-	(+)	(+)	(+)	+	++	+	+++
Metam.	(-)	+++	+++	+++	(+)	(-)	(-)	(+)	(-)	(-)
Lucan	-	(-)	+++	+++	--	(-)	--	(-)	(-)	(-)
Persius	(-)	--	---	---	(-)	(+)	(+)	(+)	(+)	(+)
Silvae	(-)	(+)	+++	+++	(-)	(-)	(-)	(-)	(-)	(-)
Thebais	-	(+)	+++	+++	(+)	(-)	(-)	(+)	(-)	(-)
Arg.	(-)	(+)	+++	+++	(-)	(-)	(-)	(+)	(-)	(-)
Juvenal	---	(+)	---	---	(-)	(+)	(-)	(-)	(+)	(-)

to the relative frequencies, and finally we compute Spearman's rank correlation coefficient. In this way we obtain the following coefficients:

No Punctuation: $r_s = 0.2889$

Strong Punctuation: $r_s = 0.5104$

Weak Punctuation: $r_s = 0.4661$

The corresponding two-tailed probabilities are $P > 0.20$, $P < 0.05$ and $P < 0.10$. Thus there is, at least in the case of strong punctuation, a positive relation between the frequencies of liaisons CV and VC, although in the case of weak punctuation the above hypothesis might be confirmed as well if we increase the number of authors/works studied.

As a final point in the analysis of GÉRARD's Tables 3 and 4 we now test whether the frequencies in GÉRARD's Table 4 are related to the chronological order of the authors/works. We proceed in the same way as with GÉRARD's Table 1 and get the following rank correlation coefficients (the corresponding two-tailed probabilities are given in brackets):

Elision: $r_s = -0.5529$ $(P < 0.05)$

CV: $r_s = 0.5361$ $(P < 0.05)$

VC: $r_s = 0.4853$ $(P < 0.10)$

Total: $r_s = 0.6059$ $(P < 0.02)$

As the probabilities indicate, in all cases but one the correlation between 'chronological order of authors/works' and 'order of relative frequencies' is significant on at least the 5% level. We thus may conclude that in trochaic divisions not marked by punctuation the use of liaison CV and probably also of liaisons VC tends to increase from Ennius to Juvenal, whereas the use of elisions tends to decrease.

4. Analysis of GÉRARD's Tables 5, 6 and 7

Since in GÉRARD's Tables 5, 6, and 7 the liaisons VC are in almost all authors/works more frequent than the liaisons CV in the category 'content word (*mots de sens plein*)' and less frequent than the liaisons CV in the category 'function word (*mots-outils*)', I will not analyse the relation between the variables 'kind of liaison' and 'kind of word' separately for each author/work but compare the different totals in GÉRARD's Tables 5, 6 and 7 directly. We again use the χ^2 test without taking the number of verses analysed into account.

The test on the totals of GÉRARD's Table 5 as represented in the contingency table above GÉRARD's Table 5 yields $\chi^2 = 122.03$ with df = 1, which corresponds to a φ coefficient of 0.4440. Analogously we obtain for GÉRARD's Table 6 $\chi^2 = 51.79$ and $\varphi = 0.2821$ and for GÉRARD's Table 7 $\chi^2 = 116.93$ and $\varphi = 0.5227$. As could be expected, the χ^2 values are in all three cases highly significant. Thus the hypothesis of an association between the variables 'kind of liaison' and 'kind of word' is confirmed. As is however indicated by the φ coefficients, the strength of this association seems to be greater in the case of trochaic divisions marked by

weak punctuation than in that of trochaic divisions marked by strong punctuation. It should be noted that this statement is based on the analysis of 68698 verses, whereas in the case of GÉRARD's Tables 1 to 5 our conclusions are based on the analysis of 13920 verses.

5. Conclusion

It would be possible, indeed desirable, to continue analysing GÉRARD's tables in the way illustrated above. But I hope that the analyses carried out thus far in this short note have demonstrated clearly enough how to apply the procedures presented to data from GÉRARD which we have not yet analysed. Furthermore I trust that this short illustration of the use of inferential statistics has been convincing enough to show the advantages of applying inferential statistics to data such as GÉRARD presents.

References

BISHOP, Y.M.M., FIENBERG, S.E., HOLLAND, P.W., Discrete Multivariate Analysis: Theory and Practice. Cambridge, Mass., The M.I.T. Press 1975

HABERMAN, S.J., The Analysis of Residuals in Cross-Classified Tables. Biometrics 29, 1973, 205-220

ŘEHÁK, J., ŘEHÁKOVÁ, B., Analyse von Kontingenztafeln: Zwei Grundtypen von Aufgaben und das Vorzeichenschema. In: MATTHÄUS, W. (Ed.), Glottometrika 3. Bochum, Brockmeyer 1980, 1-29

REYNOLDS, H.T., The Analysis of Cross-Classifications. Riverside, N.J., The Free Press 1977

Annotated Bibliography on the Statistical Study of Hexameter Verse

Ulrike Job, Bochum

In the context of the present volume a bibliograhy on studies of hexameter verse should preferably focus on the quantitative aspects of these studies. One reason for this constraint lies in the fact that the dilemma of philological interpretations of isolated metrical and textual phenomena - often open to serious doubt with regard to their objectivity - can only be diminished by basing the analyses on larger corpora, large enough to allow us to set up well-founded hypotheses and/or to test hypotheses already established. This latter aim, i.e. testability, is an important claim in epistemology; untestable hypotheses are characteristic of what BUNGE (1967: 36f.) calls 'pseudoscience.'

The corpus of hexameter poetry clearly is large enough and at the same time sufficiently limited to make the application of statistical methods useful. In addition, a diagnostic function for quantitative analyses may be found in tests of the homogeneity of texts with regard to their assumed authorship.

For this reason, the present bibliography includes only publications on hexameter verse containing numerical data on metrical and textual details of the hexameter. These data are given either in absolute frequencies or in percentages with an indication of the sample size and/or population involved. Items without statistically relevant numerical data, for instance containing percentages without reference to a population, have not been included. However, works dealing with statistical methods and their application to hexameter (cf. entries 3, 19, 67, 82, 83, 93-97, 147, 159, 170, 171, 236, and 239) have been quoted, even if they do not contain a considerable amount of empirical data.

The period under review covers the time from 1866 (entries 44 and 145) to 1981. An important landmark is the introduction

of the computer in the early sixties (cf., e.g., entry 31), by which it became possible to handle larger and more complex corpora than before. Parallel to this, a more explicit and better founded use of statistical methods was developed.

The items in this bibliography have been supplied with short annotations referring to the nature and sources of the numerical data and/or the statistical methods involved. These annotations are as a rule the result of personal examination of the works quoted. Only occasionally have annotations been made on the basis of personal communications. Items inaccessible to me have been included only if there was sufficient reason to believe that they contain numerical data on hexameter (e.g. for part of the works published in *Pallas* that I did not have an opportunity to examine). All these items, i.e., those which I could not examine myself or which were not supplied with an annotation based on the author's information, have been marked with an asterisk (*). Since only the most common Latin and Greek works and authors' names have English equivalents, I have in all cases used the originals (Ovidius, Livius, Odysseia, etc.).

An enterprise like this would not have been possible without the assistance and criticism of scholars who are known specialists in this field of study. I wish to thank all those who have contributed to the completeness of this bibliography by kindly answering my letters or by commenting on an earlier draft of the list of items, especially G. Altmann, D.L. Clayman, J. Cupaiuolo, S.I. Gindin, N.A. Greenberg, A. Kleinlogel, H. Kreuzer, M.R. Mayenowa, A.Q. Morton, W. Ott, J. Põldmäe, S. Prete, L. Stephens, P. Tordeur and St.V.F. Waite. Special mention must be made of E. Liénard, R. Schmiel, and J. Soubiran, who made most helpful comments on a preliminary version of this bibliography, and whose assistance far exceeded that which I could have ever expected to receive. It should be clear, however, that the responsibility for any shortcomings is entirely my own.

Reference:

BUNGE, M., Scientific Research I. The Search for system. Berlin, Heidelberg, New York: Springer 1967.

BIBLIOGRAPHY

List of Abbreviations

Aevum	Aevum. Rassegna di Scienze storiche, linguistiche e filologiche, pubblicata a cura della Facoltà di Lettere e Filosofia dell' Università Cattolica del Sacro Cuore.
AJPh	American Journal of Philology.
CPh	Classical Philology.
CQ	The Classical Quarterly.
Em	Emérita. Boletín de linguística y filología clásica.
Glotta	Glotta. Zeitschrift für griechische und lateinische Sprache.
Hermes	Hermes. Zeitschrift für klassische Philologie.
HSPh	Harvard Studies in Classical Philology.
KZ	Zeitschrift für vergleichende Sprachforschung auf dem Gebiet der indogermanischen Sprachen, begründet von A. Kuhn.
Latomus	Latomus. Revue d'études latines.
LD	Linguistik und Didaktik.
Lg	Language. Journal of the Linguistic Society of America.
LiLi	LiLi. Zeitschrift für Literaturwissenschaft und Linguistik.
Mind	Mind. A Quarterly Review of Psychology and Philosophy.
Mn	Mnemosyne. Bibliotheca Philologica Batava.
MPh	Modern Philology.
Pallas	Pallas, fasc. 3 des Annales de l'Université de Toulouse-Le Mirail.
Phoenix	The Phoenix. The Journal of the Classical Association of Canada.

PhQ	Philological Quarterly.
REL	Revue des Etudes Latines.
RELO	Revue de l'organisation internationale pour l'étude des langues anciennes par ordinateur.
RhM	Rheinisches Museum für Philologie. Neue Folge.
RSC	Rivista di Studi Classici.
SbSAW	Sitzungsberichte der Sächsischen Akademie der Wissenschaften zu Leipzig, Philologisch-historische Klasse.
TAPA	Transactions and Proceedings of the American Philological Association.
TPhS	Transactions of the Philological Society.
UCPL	University of California Publications in Linguistics.
Vergilius	Vergilius. The Vergilian Society of America.
WS	Wiener Studien. Zeitschrift für klassische Philologie und Patristik.
ZPhon	Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung.

1. ALLEN, F. (1879), Ueber den Ursprung des homerischen Versmasses. KZ 24, 556-592.

   Data on trochaic caesura in Homeros.

2. ALLEN, W.S. (1966), Prosody and prosodies in Greek. TPhS, 107-148.

   Assumes (excluding the position in pausa)
   a) that words were primarily stressed on their last heavy syllable and
   b) that a secondary stress fell on the preceding heavy syllable if separated from the primary stress by at least one mora of quantity;
   data on the first 100 lines of Sophokles' Antigone and Homeros' Odysseia; frequencies of poly-syllabic word patterns.

3. ALTMANN, G. (1981), The homogeneity of metric patterns in hexameter. R. Grotjahn (ed.): Hexameter studies. Bochum: Brockmeyer, 137-150. (= Quantitative linguistics 11).

   Method to determine with the help of information statistics whether texts or groups of texts can be considered as homogeneous; data on several Latin and German authors.

4. Arbeitsgruppe für lateinische Metrik und Stilistik, Tübingen (Hildegard Cancik-Lindemaier; Hubert Cancik; Markus Hen-

rich; Dirk Kottke; Wilhelm Ott; Reinhard Vogel) (1972), Zur Elision anapästischer Wörter bei Vergil und Statius. Glotta 50, 97-120.

Frequencies of different types of anapaestic elision; computational analysis.

5. ARTYMOWICZ, A. (1909), Der Wechsel von et und que zu Beginn lateinischer daktylischer Verse vor Ennius bis Corippus. WS 31, 38-81.

Tendencies concerning the change of et and que at the beginning of the line.

6. ATHANASSAKIS, A. (1970), Hiatus, word end, and correption in Hesiod. Hermes 98,2, 129-145.

Data on long and short vowels in the case of hiatus; manual analysis.

7. * BAILEY, J. (1975), Toward a statistical analysis of English verse. Lisse: Peter de Ridder (= PdR press publication in English metrics 1).

8. * BARLOW, J.F. (1956), Metrical word-types in the Latin hexameter. Susquehanna Univ. studies 5, fasc. 4, 232-262.

9. BEEKES, R.S.P. (1972), On the structure of the Greek hexameter. 'O'Neill' interpreted. Glotta 50, 1-10.

Discussion of O'Neill's data (O'Neill 1942) and derivation of 6 rules based on these data concerning the Greek hexameter.

10. BERSCHIN, H. (1970), Automatische metrische Analyse lateinischer Verse. LD 1, 72-80.

Computational analysis of the number of syllables, words, and synaloephae in the Latin hexameter.

11. BOLDRINI, M. (1948), Esametri. Contributi del laboratorio di statistica, serie sesta, dell'Università Cattolica del Sacro Cuore 21, 76-79.

15 different writers; discussion of Drobisch's data (Drobisch 1866); distribution of d/s in the first four feet.

12. BONARIA, M. (1962), L'elisione parziale nella poesia latina. Hermes 90, 173-186.

Data on the correptio epica in Homeros, Hesiodos, and others.

13. BRÉGUET, E. (1946), Le roman de Sulpicia. (thèse) Genève.

Distribution of d/s, proportions (%) of dactyls in the first four feet of the hexameter and incisions within the hexameter (absolute numbers and percentages); data on Sulpicia (Elegiae 4, 2-6), Tibullus (Elegiarum libri 1,1 and 2), Propertius (Elegiarum libri 1-4), Ovidius (Amores, Heroides, Remedia amoris, Fasti).

14. CARPENTER, A.H. (1930), Hypermetric lines and interlinear hiatus in Latin hexameter verse. PhQ 9, 351-362.

Study of interlinear hiatus in the hexameter poetry of Catullus, Lucretius, Cicero, Vergilius, and Horatius Flaccus.

15. CARTAULT, A. (1911), La distique élégiaque chez Tibulle, Sulpicia, Lygdamus. Université de Paris (= Bibliothèque de la Faculté des Lettres 27).

Numerical information about the prosodic structure, elisions, and caesurae in Tibullus (Elegiarum libri without 4, 2-6 and 13-14), Sulpicia (short works), and Lygdamus (Elegiae 1-6).

16. CAVALLIN, C. (1896), De caesuris quarti et quinti trochaeorum hexametri apud Latinos poetas coniunctis. Lund: Norcopiae.

Data on Ennius, Lucilius, Lucretius, Cicero, Catullus, Vergilius, Horatius, Tibullus, Propertius, Ovidius, Germanicus, Manilius, Grattius, Calpurnius, Lucanus, Persius, Valerius Flaccus, Silius Italicus, Iuvenalis, Martialis, et al.

17. CHRISTENSEN, H. (1908), -que -que bei den römischen Hexametrikern (bis etwa 500 n. Chr.). Archiv für lateinische Lexicographie und Grammatik 15, 165-211.

Frequencies of trisyllables after the penthemimeral and hephthemimeral caesurae in various Latin poets from Ennius to Martialis.

18. CLARKE, W.M. (1972), Intentional rhyme in Vergil and Ovid. TAPA 103, 49-77.

A manual study of the Aen., Metam., and samples from other authors which indicates that rhyme, often not syntactically dictated, occurs frequently near major sense pauses in Vergilius and Ovidius.

19. CLAYMAN, D.L. (1981), Sentence length in Greek hexameter poetry. R. GROTJAHN (ed.): Hexameter studies. Bochum: Brockmeyer, 107-136. (= Quantitative linguistics 11).

Conclusion that sentence length distribution measured by words differs from sentence length distribution measured by phonemes (or by syllables); data on 23,049 sentences gathered from all hexameter poetry from the eigth to the second centuries B.C.

20. CLAYMAN, D.L. / NORTWICK, Th. van (1978), Enjambement in Greek hexameter poetry. TAPA 107, 85-92.

Summary of Parry's statistics (Parry 1929), who counted the occurrences of enjambement in 600-line samples of each of the Homeric poems, the Argonautica of Apollonios and the Aeneis of Vergilius. To correct Parry's unreliable data, the author provides new data on enjambement.

21. CONRAD, C.W. (1965), Traditional patterns of word-order in Latin epic from Ennius to Vergil. HSPh 69, 195-258.

Patterns of Sperrung in the Latin hexameter of Ennius, Cicero, Lucretius, Catullus, and Vergilius.

22. COSTA, T. (1957), Formele hexametrului la Ovidiu Publius Ovidius Naso. Biblioteca Antică Studii 2, 211-332. Bucureşti: Editura Academiei Republicii Populare Romîne.

Distribution of d/s in the works of Ovidius and the Ilias of Homeros.

23. CRISI, V. (1938), De re metrica et prosodica D. Magni Ausonii I: De hexametris et pentametris. Utini: del Bianco.

Data on the structure of the hexameter of Ausonius.

24. CUPAIUOLO, F. (1963), Un capitolo sull'esametro latino. Parole e finali dattiliche o spondaiche. Napoli: Libreria scientifica editrice (= Collana di Studi Latini 5).

Frequency of d/s in single feet in Ennius, Catullus, Lucretius, Vergilius, Horatius, Ovidius, Persius, Lucanus, Valerius Flaccus, Silius Italicus, Statius, Iuvenalis, and Claudianus.

25. CUPAIUOLO, F. (1965a), Parole giambiche nell'esametro latino. RSC 13, 31-43.

Frequencies of d/s in single feet; data on various Latin authors from Ennius to Claudianus.

26. CUPAIUOLO, F. (1965b), Studi sull'esametro di Catullo. Napoli: Libreria scientifica editrice (= Collana di Studi Latini 10).

Analysis of the Carmina docta of Catullus: Count of final words as to number of syllables; d/s-types in Carmen LXIV with raw data and percentages which are compared with absolute numbers and percentages in Ennius, Lucilius, and Lucretius; table of caesurae in Carmen LXIV; absolute numbers and percentages of different metrical word types; initial word types in absolute numbers and percentages in different Carmina.

27. CUPAIUOLO, F. (1971), Sul ricorrere nell'esametro latino di parole con la forma prosodica di pirrichio. Bolletino di Studi Latini 1, fasc. 2, 240-250.

Frequencies (absolute numbers and percentages) of pyrrhic words in Vergilius' Bucolica; the position with lower frequencies of pyrrhic words are the fourth and the first foot.

28. CUTT, T. (1936), Meter and diction in Catullus' hendecasyllabics. (diss.) Chicago.

Attempt to estimate the influence of meter on diction in the hendecasyllabics of Catullus by examining the relationship between the requirements of the verse and the language used; tables of different verse-types and of the absolute number of each word-type used.

29. DARQUENNE, M. (1970), Essai sur la place du nom et de l'adjectif dans l'hexamètre latin. Bruxelles: Université Libre de Bruxelles.

Positions of noun and adjective in the hexameter of Ovidius, third book of Metamorphoseon libri (733 verses), Statius, first book of Thebais (720 verses), and Sidonius Apollinarius, Carmen V (603 verses); metrical patterns of nouns and adjectives.

30. DAVIS, G.T. (1959), A study in Latin prosody. Short final vowel before initial consonant groups composed of stop and liquid. (diss.) University of Pennsylvania, microfiche 59-4610 (= Language and Literature Classical).

Presentation of the percentage of short final vowels before all initial consonants other than stop-liquid combinations and presentation of the percentage of short final vowels before stop-liquid combinations in Ennius, Catullus, Ovidius, Tibullus, Horatius, Vergilius, and others.

31. DELATTE, L./ EVRARD, E. / GOVAERTS, S. / HAZETTE, P. (1962), Sénèque: Consolation à Polybe: Index verborum. Relevés statistiques. Liège: L.A.S.L.A. (= Travaux publiés par le laboratoire d'analyse statistique des langues anciennes 1).

Index verborum; word frequencies; distribution of words according to their grammatical categories; computational analysis.

32. DELATTE, L. / EVRARD, E. / GOVAERTS, S. / HAZETTE, P. (1963), Sénèque: Consolation à Helvia: Index verborum. Relevés statistiques. La Haye: Mouton (= Travaux publiés par le laboratoire d'analyse statistique des langues anciennes 2).

Computational analysis of word frequencies, of the distribution of words according to their grammatical categories, and of the position of words and sentences.

33. DELATTE, L. / EVRARD, E. / GOVAERTS, S. (1964), Sénèque - Consolation à Marcia. Index Verborum. Relevés statistiques. La Haye: Mouton (= Travaux publiés par le laboratoire d'analyse statistique des langues anciennes 3).

Vocabulary of the Ad Marciam de Consolatione; frequency distribution according to grammatical categories; phrase length; word order.

34. * DELATTE, L. / EVRARD, E. (1966), Sénèque, De constantia sapientis. Index verborum. Relevés statistiques. Avec la collaboration de A. Bodson et S. Govaerts. La Haye: Mouton (= Travaux publiés par le laboratoire d'analyse statistique des langues anciennes 7).

35. DEVINE, A.M. / STEPHENS, L. (1976), The Homeric hexameter and a basic principle of metrical theory. CPh 71, 141-163.

Statistics on the distribution of syllable types in Homeros compared with a sample of prose from Thukydides. Construction of a probabilistic model demonstrates that the frequency distributions of VC, CV, and CVC syllable-types in longum and breve positions is an automatic result of fitting Greek linguistic structure to non-durationally motivated aspects of hexameter structure and not evidence for recent theories of a multiplicity of metrical elements corresponding to postulated differences in phonetic duration within the phonological category of the heavy syllable.

36. DOBRZYŃSKA, T./KOPCZYŃSKA, Z. (1979), Tonizm. Poetyka. Zarys encyklopedyczny. Dział 3. Wersyfikacja. Tom 5. Wrocław: Ossolineum.

Presentation of the syllabic and accentual make-up of different verses with regard to rhythm and accent, including the hexameters of Adam Mickiewicz; percentages of rhythmical types, subdivided as to number of syllables.

37. DREWITT, J.A.J. (1908), Some differences between speech-scansion and narrative-scansion in Homeric verse. CQ 2, 94-109.

Data on occurrences of break, elisions, shortenings, and overlengths in certain positions, tabulated according to narration and speech in the Ilias and the Odysseia of Homeros.

38. DREXLER, H. (1951), Hexameterstudien. Aevum 25, 435-466 and 512-547.

Observation that in verses divided into three parts by the semiternaria and semiseptenaria after iambic words, the syntactical break is placed as a rule before the semiternaria, whereas especially spondaic words are found in a close syntactical relationship or in Sperrung with others; data from the first 1000 verses each of Lucanus, Ovidius (Metam.), Vergilius (Aen.), Horatius (Sat. I), and Lucretius, also verses of Catullus, Lucilius, and Ennius.

39. * DREXLER, H. (1952), Hexameterstudien V. Verse ohne semiquinaria oder mit überwiegender semiseptenaria. Em 20, 427-460 (to be continued).

40. * DREXLER, H. (1953a), Hexameterstudien V. Em 21, 162-176.

Continuation of Drexler 1952.

41. DREXLER, H. (1953b), Hexameterstudien. (II-III). Iambische und spondeische Wörter vor der Semiquinaria. Verse mit überwiegender Semiquinaria. Wortschluss nach der vierten Hebung oder der vierten Senkung. Salamanca: C.S.I.C. (= Theses et studia philologica Salamanticensia 5).

Assumption that in verses with prevailing semiquinaria the syntactical break is placed before this caesura; data as in Drexler 1951.

42. DREXLER, H. (1953c), Hexameterstudien IV. Der Bau der beiden Vershälften. Göttingen: Vandenhoeck & Rupprecht, Microfilm.

Distribution of metrical word types before caesura; monosyllables after the semiquinaria; spondaic or dactylic word ends after the 4th stressed or 4th unstressed syllable respectively; data on Lucanus, Ovidius, Vergilius, Horatius, Catullus, Lucretius, Lucilius and Ennius.

43. DREXLER, H. (1956), Hexameterstudien, V-VI. Madrid: C.S.I.C. (= Manuales y anejos de Em 14).

(Pages 5-52 have already been published in Emerita 20, 1952, 427-460, cf. Drexler 1952); numerical information about verses with semiquinaria as main caesura or with prevailing semiseptenaria; frequency of synaloepha in single verses and in the work as a whole; frequency of monosyllables before caesura (without synaloepha); data as in Drexler 1951.

44. DROBISCH, M.V. (1866), Ein statistischer Versuch über die Formen des Lateinischen Hexameters. SbSAW 18, 75-139.

Distribution of d/s in several Roman authors; the hexameter of Vergilius is regarded as the classical hexameter in which the ratio of d/s is 4 : 5.

45. DROBISCH, M.V. (1868a), Weitere Untersuchungen über die Formen des Hexameters des Vergil, Horaz und Homers. SbSAW 20, 16-65.

The distribution of d/s is not deliberate, but determined by principles which are followed by the author unconsciously; data from Vergilius' Georgica, Aeneis, Bucolica, Horatius' Satirae, Epistulae, and Homeros' Ilias, Odysseia.

46. DROBISCH, M.V. (1868b), Über die Formen des deutschen Hexameters bei Klopstock, Voss und Goethe. SbSAW 20, 138-160.

Data on d/s in Klopstock's Messias, in Voss' Luise, and Goethe's Reinecke Fuchs and Hermann und Dorothea; comparison with data on the Latin and Greek hexameter.

47. DUBOIS, Ch. (1935a), Lucrèce, poète dactylique. Strasbourg: Librairie universitaire d'Alsace.

Frequency of d/s in Lucretius, book 1-6; symmetry of words, frequency of single words; constitution of the first two feet; placement of the hiatus (comparison with other authors); placement of the aphaeresis; occurrence of elision.

48. DUBOIS, Ch. (1935b), La métrique de Lucrèce comparée à celle de ses prédécesseurs Ennius et Lucilius. (thèse) Lille, Strasbourg.

Metrical analysis comparing for example data on the repartition of d/s, on penthemimeral, hephthemimeral, and trihemimeral caesura, on word dispositions, on the ends of the hexameter line, and on coincidences of feet and words in Ennius, Lucilius, and Lucretius.

49. DUCKWORTH, G.E. (1962), Structural patterns and proportions in Vergil's Aeneid. A study in mathematical composition. Ann Arbor: The University of Michigan press.

Discussion of the golden mean ratio in the Georgica and the Bucolica, comparison of Vergilius with Catullus (Carmina 64), Lucretius (De rerum natura 1), Horatius (Satirae, Epistulae, Ars poetica), and Maphaeus Vegius.

50. DUCKWORTH, G.E. (1964), Variety and repetition in Vergil's hexameters. TAPA 95, 9-65.

Distribution of d/s patterns in the Aeneis, Georgica, and Bucolica of Vergilius; some indications referring to the hexameter of Homeros and Vergilius.

51. DUCKWORTH, G.E. (1965), Horace's hexameters and the date of the Ars Poetica. TAPA 96, 73-95.

Variation and repetition, opposite and reverse patterns in Horatius.

52. DUCKWORTH, G.E. (1965-66), Hexameter patterns in Vergil. Proceedings of the Vergil society 5, 39-49.

The patterns that Vergilius favors in the Aeneis (10-12), his use of repeated, opposite, and reverse patterns in adjacent lines, the metrical differences between Aeneis (10-12) and the Georgica and Aeneis (1-9), the favorization of dsss is affected by the style; the more frequent a pattern, the less frequent its opposite in the Aeneis.

53. DUCKWORTH, G.E. (1966a), The non-Ovidian nature of the Halieutica. Latomus 25, 756-768.

Distribution of d/s patterns in the first four feet in the hexameter of Ennius, Lucilius, Cicero, Catullus, Lucretius, Vergilius, and Horatius; also percentages on homodyne with these authors.

54. DUCKWORTH, G.E. (1966b), Studies in Latin hexameter poetry. TAPA 97, 67-113.

Continuation of the analyses in Duckworth 1966a.

55. DUCKWORTH, G.E. (1966c), Vergil's subjective style and its relation to meter. Vergilius 12, 1-10.

Correlation of meter and style,
test results: - the most frequent pattern is dsss in the first four feet,
- emotional scenes have a low proportion of dsss patterns;
data from Aeneis and Georgica.

56. DUCKWORTH, G.E. (1967), Five centuries of Latin hexameter poetry: Silver Age and late empire. TAPA 98, 77-150.

Distribution of d/s of Latin authors from the Silver Age up to the Christian post-Ovidian authors.

57. DUCKWORTH, G.E. (1968), A rare type of first foot dactyl (three words). AJPh 89, 437-448.

Frequency of three word dactyls in the Latin hexameter (from Ennius and Lucilius to Arator and Corippus). The list gives the totals (first foot only) for the poets or poems and the relative frequencies. For final summary of the frequency of a first foot dactyl composed of three words a table is presented with totals for the various literary periods.

58. DUCKWORTH, G.E. (1969a), Maphaeus Vegius and Vergil's Aeneid: A metrical comparison. CPh 64, 1-6.

Comparison of the distribution of d/s in Vegius and Vergilius.

59. DUCKWORTH, G.E. (1969b), Vergil and classical hexameter poetry. A study in metrical variety. Ann Arbor: The University of Michigan press.

Data range from Ennius to the time of Justinian; distribution of d/s.

60. DUCKWORTH, G.E. (1972), Milton's hexameter patterns: Vergilian or Ovidian? AJPh 93,1, 52-60.

Frequency of d/s in Milton, Vergilius, and Ovidius suggests affinity of Milton's hexameter to both authors' practices.

61. DYER, R.R. (1967), Toward computational procedures in Homeric scholarship. RELO, fasc. 4, 1-54.

Application of an automatic scansion program to the last book of the Odysseia; grammatical analysis of the metrical word shape and examination of compositional word-type.

62. ESKUCHE, G. (1890), Die Elisionen in den zwei letzten Füssen des lateinischen Hexameters, von Ennius bis Walahfridus Strabo. RhM 45, 236-264 and 385-418.

Numerical information about elided short and long vowels, elided internal syllables, and about elisions in general; data on Ennius, Lucilius, Cicero, Lucretius, Valerius Cato, Catullus, Vergilius, Horatius, Tibullus, Lygdamus, Sulpicia, Propertius, and Walahfridus Strabo.

63. EVRARD, E. (1972), Scansion automatique de l'hexamètre grec. RELO, fasc. 4, 1-33.

Examination of caesura; analysis of the distribution of d/s in five Idylls (I, VII, XIII, XV, XVI) of Theokritos.

64. EVRARD, E. (1978), Quelques observations sur la 5^{e} Bucolique de Virgile. Les Etudes Classiques 46,4, 327-338.

Comparison of metrical patterns in Mopsus' and Menalcus' presentations, there are certain statistically significant differences in patterns at ends of lines.

65. FORTMANN, A. (1909), Quaestiones in Lucanum metricae. (diss. in.) Greifswald.

Numerical information (absolute numbers and percentages) about proportions of d/s, the penthemimeral and the hephthemimeral caesurae, and about different sorts of elisions in Vergilius (Aeneis 1, 4, 8) and Lucanus (1-10).

66. FRANK, E. (1970), The ictus in the Vergilian and silver epic hexameter. Istituto Lombardo. Accademia di Scienze e Lettere. Rendiconti. Classe di Lettere e Scienze Morali e Storiche 104, 327-331.

Percentages of heterodyne structures for each of the first four feet and combined percentages for all the first four feet in Vergilius and the epic poets of the Silver Age, in addition, the corresponding frequencies in Ovidius' Metamorphoseon libri (1. 1-400, 2. 1-200) for purposes of comparison; it is the aim of this study to investigate the handling of metrical stress throughout the whole line by Vergilius and the Silver Age epic poets.

67. FUCKS, W. (1955), Mathematische Analyse von Sprachelementen, Sprachstil und Sprachen. Köln/Opladen: Westdeutscher Verlag. Arbeitsgemeinschaft für Forschung des Landes NRW.

Frequencies of metrical elements in Rilke, Goethe, and Sallustius; various methods for analyzing hexameter verse.

68. GARSON, R.W. (1968), Metrical statistics of Valerius Flaccus' Argonautica. CQ N.S. 18,2, 376-379.

Distribution of d/s; elisions; results of manual analysis.

69. GASPAROV, M.L. (1975), Prodrom, Cec i nacional'nye formy gekzametra. L.A. Frejberg (ed.): Antičnost' i Vizantija, Moskva: Nauka, 362-385.

Frequencies of spondees in metrical feet; analysis of Greek, Latin, Byzantine, German, English, Polish, and Russian hexameter; hypotheses about national forms of hexameter verse.

70. GEORGACAS, D.J. (1948), The causes of vocalic and consonantal aphaeresis and prothesis in late and modern Greek. TAPA 79, 340.

Abstract of an inaccessible monograph; quantitative information on consonant and vowel patterns in Homeric Greek and Modern Greek.

71. GÉRARD, J. (1972), La ponctuation trochaïque dans l'hexamètre latin d'Ennius à Juvénal. Exemplaires dactylographiés déposés à la Bibliothèque Universitaire de Paris-Sorbonne sous la cote W 1972 (10 bis) 1-2, 4^{o}, Paris.

Quantitative and statistical analysis concerning the division of short syllables of the dactyl between finals and initials of words which are marked by weak or strong rhetorical punctuation at different positions in the hexameter; data from Ennius, Lucilius, Catullus, Lucretius, Vergilius (Buc., Georg., Aen.), Horatius (Sat., Epist.), Ovidius (Met.), Lucanus, Statius (Silv., Theb.), Persius, Valerius Flaccus, and Iuvenalis.

72. GÉRARD, J. (1980), La ponctuation trochaïque dans l'hexamètre latin d'Ennius à Juvénal. Recherches sur les modalités de l'accord entre la phrase et le vers. Paris: Les Belles Lettres (= Collection d'études latines 36).

Analysis as in Gérard 1972.

73. GÉRARD, J. (1981), Liaisons syllabiques dans les partages trochaïques de l'hexamètre latin. R. GROTJAHN (ed.): Hexameter studies. Bochum: Brockmeyer, 168-211 (= Quantitative linguistics 11).

On the basis of the differentiation between several types both of trochaic and syllabic liaison the conclusion is drawn that vowel-consonant liaisons tend to be more frequent than consonant-vowel liaisons in the category 'content word' and less frequent than consonant-vowel liaisons in the category 'function word'; data on several Latin authors from Ennius to Iuvenalis.

74. GIARRATANO, C. (1908), De M. Val. Martialis re metrica. Napoli: Detken et Racholl.

Frequencies of d/s in single feet (absolute frequencies and percentages); proportions of monosyllabic, bisyllabic, trisyllabic etc. words (absolute numbers and percentages); elisions.

75. GLENN, J.M. (1939), The neuter plural in Latin iambic and trochaic verse. Lg 15 Suppl. (= Lg Diss. 30).

Frequencies of nominative, accusative, and vocative according to gender and number in Plautus and Terentius.

76. GODOLPHIN, F.R.B. (1939-40), Notes on the technique of Milton's Latin elegies. MPh 37, 351-356.

Frequency of different metrical devices in Tibullus, Propertius, Ovidius, and Milton; similarity in technique between Milton and the Roman elegists (distichs, percentage of rhymed hexameter).

77. GÖTZINGER, M. (1869), Zum deutschen Hexameter. Neue Jahrbücher für Philologie und Paedagogik 100, (Abt. 2, Jg. 15), 145-151.

Proportion of d/s in single feet of different German authors.

78. GOODWIN, W.W. (1912), Syntax of the moods and tenses of the Greek verb. London: Macmillan and Co (2nd edition).

Statistics on the use of the final particle in pure final clauses from fourteen Greek authors, especially from the Ilias and the Odysseia (Appendix III).

79. GOVAERTS, S. (1966), Les initiales de vers chez Lucrèce et Virgile. Statistique et analyse linguistique. Colloque de Strasbourg 20-22 avril 1964, Paris: Presses Universitaires de France, 41-49.

Study of the repetition rate of verse-initial phonemes in Lucretius (De rerum natura) and in Vergilius (Aeneis): the preference for certain gaps between specified verse-initials could have phonetic or semantic grounds.

80. GOVAERTS, S. (1978), De l'utilisation des index et des relevés grammaticaux et syntaxiques. RELO, fasc. 4, 37-76.

Description of morphologic, syntactical, and lexicographic analyses, with examples of indices, tables, and graphs from various Latin authors.

81. GREENBERG, N.A. (1967a), Vergil and the computer: fourth foot texture in Aeneid I. RELO, fasc. 1, 1-16.

Testing of some of W.F.J. Knight's speculations in his Accentual symmetry in Vergil (CR 54, 1940, 93-95); tests show that Vergilius' sequence of homodyned and heterodyned fourth feet in the Aeneis is not random.

82. GREENBERG, N.A. (1967b), Scansion purement automatique de l'hexamètre dactylique. RELO, fasc. 3, 1-30.

Description of his scansion program including its shortcomings.

83. GREENBERG, N.A. (1970a), Words and syllables: Four Eclogues. RELO, fasc. 2, 5-49.

Discussion of various statistical procedures applied to Bucolica (1, 4, 6, and 9) of Vergilius; comparison of the distribution of words according to the number of syllables they contain, and according to the number of morae they contain.

84. GREENBERG, N.A. (1970b), Metrical expectations in the Ars Poetica. RELO, fasc. 3, 111-129.

Frequencies of different line types and pairs of word junctures are not statistically significant, except that monosyllables tend to occur at the beginning of the line; data from Horatius.

85. GREENBERG, N.A. (1970c), Words and syllables: Four Eclogues: A correction. RELO, fasc. 3, 131

Addition of some observations to the article of the same title in RELO 1970, 2, 5-49.

86. GREENBERG, N.A. (1970d), The hexametrical maze. RELO, fasc. 4, 17-63.

Discussion of the extent to which a word pause in one position in the verse conditions others elsewhere; lengthening of monosyllables and words repeated in the same line position; data from Vergilius, Bucolica and Aeneis I, Horatius, Ars Poetica, Ovidius, Metamorphoseon libri.

87. GREENBERG, N.A. (1972a), Epanastrophe in Latin poetry. RELO, fasc. 2, 1-17.

The hexameter verse of Lucretius lacks the suppleness and flexibility which can be demonstrated quantitatively in the verse of Vergilius and Ovidius; treatment of syllables and letters repeated over word junctures and of letter combinations at word boundaries distinguishes Lucretius from Catullus, Vergilius, Ovidius, Lucanus, Statius, and Sallustius.

88. GREENBERG, N.A. (1972b), Line-initials in the Georgics. RELO, fasc. 4, 55-58.

Supports the hypothesis that the distribution of initial letters in lines is random by developing the data given by G. Herdan in the Calculus of linguistic observations (1962).

89. GREENBERG, N.A. (1975), Line-initials in Latin hexameter. RELO, fasc. 1, 1-25.

Examination of the intervals between lines beginning with the same letter in selected poets, presented in 14 tables; conclusion that there is insufficient evidence to reject the hypothesis that poets pay no deliberate attention to the initial letters of lines.

90. GREENBERG, N.A. (1978), Metrical shape, initial stress, and cross-tabulation. RELO, fasc. 3, 1-44.

Statistical analysis of several specimens representative of Latin hexameter poetry. Reports on combinations of dactyls and spondees and, tentatively, on series of feet with stress on the first syllable in Latin hexameter poetry.

91. GREENBERG, N.A. (1981), Vocalic initials in Aeneid 12. R. GROTJAHN (ed.): Hexameter studies. Bochum: Brockmeyer, 151-167 (= Quantitative linguistics 11).

Distribution of initials and elisions at each position in the verse line. In position 3, 6, and 9 the proportion of vocalic initials is considerably higher than at the beginning of the verse, which seems to be true of Latin hexameter verse in general.

92. GROOT, A.W. de (1935), Wesen und Gesetze der Cäsur. Mn 2, 81-154.

Positions of word boundaries and of syntactical boundaries in Homeros (Ilias 1, 644 lines) and Vergilius (Aeneis 6, 1-397).

93. GROTJAHN, R. (1979), Linguistische und statistische Methoden in Metrik und Textwissenschaft. Bochum: Brockmeyer (= Quantitative linguistics 2).

Presentation of the model of a non-homogeneous Markov chain for analyzing Latin hexameter verses; development of probabilistic generative rules for Latin hexameters (chapter 13).

94. GROTJAHN, R. (1980), The theory of runs as an instrument for research in quantitative linguistics. R. GROTJAHN (ed.): Glottometrika 2. Bochum: Brockmeyer, 11-43.

While most statistical methods take only the frequency of the sample elements into account, but not their arrangement, the theory of runs also makes it possible to test the arrangement of sample elements for randomness. The theory of runs is presented and its use is clarified by several examples some of which are taken from hexameter verse.

95. GROTJAHN, R. (1981a), A note on Jean Gérard's paper "Liaisons syllabiques dans les partages trochaïques de l'hexamètre latin". R. GROTJAHN (ed.): Hexameter studies. Bochum: Brockmeyer, 212-225 (= Quantitative linguistics 11).

Some methods of inferential statistics are applied to Gérard's data (GÉRARD 1981) to show how these methods can be used to test Gérard's hypotheses.

96. GROTJAHN, R. (1981b), A note on Pol Tordeur's paper "Etude statistique sur l'hexamètre d'Ausone". R. GROTJAHN (ed.): Hexameter studies. Bochum: Brockmeyer, 97-106 (= Quantitative linguistics 11).

On the basis of Tordeur's data it is argued that the computation of appropriate coefficients which measure the relation between two or more variables independently from the sampling error is necessary for justified conclusions.

97. GROTJAHN, R. (1981c), A statistical model for the analysis of the coincidence of ictus and accent. R. GROTJAHN (ed.): Hexameter studies. Bochum: Brockmeyer, 33-74 (= Quantitative linguistics 11).

The hypergeometric probability distribution is discussed as a general model for the analysis of the coincidence of ictus and accent; presentation of a table permitting an objective classification of hexameter lines into homodyne, heterodyne, and neutral.

98. HACKMANN, L.F. (1940), Servius and his sources in the commentary on the Georgics. (diss.) New York: Columbia University press.

Absolute numbers of words in different poems of Vergilius (the average is 6.44 words per verse).

99. HARKNESS, A.G. (1908), The word-group accent in Latin hexameter. CPh 3, 39-58.

Some data on word-group accent (pyrrhic, iambic, monosyllabic words in the hexameter verse) in Ovidius, Metamorphoseon libri (first 6 books), Lucanus, and Vergilius (Aeneis 1-6).

100. HEITLAND, W.W. (1899), The strong hephthemimeral pause in Latin hexameter poetry. Journal of Philology 26, 1-24.

Tables of the average occurrence of the hephthemimeral caesura in Greek writers and in Latin writers; there is a dominance of this sort of pause in Latin writers.

101. HELLEGOUARC'H, J. (1964), Le monosyllabe dans l'hexamètre latin. Essai de métrique verbale. Paris: Klincksieck.

Relative frequencies of monosyllables; monosyllables and syntax, caesura, elision; position and role of -que in the Latin hexameter; parallelism between the decrease of the number of monosyllables and the increase of -que.

102. HELLEGOUARC'H, J. (1969), La ponctuation bucolique dans les Satires de Juvénal. Etude métrique et stylistique. Mélange de linguistique, de philologie, et de méthodologie de l'enseignement des langues anciennes offerts à M. René Fohalle à l'occasion de son soixante-dixième anniversaire. Gembloux: Duculot, 173-189.

Comparisons with Lucretius, Catullus, Vergilius, Ovidius, and Lucanus suggests that bucolic punctuation is used intentionally.

103. HELLEGOUARC'H, J. (1970), Les rapports de la phrase et du vers dans les constructions métriques et strophiques latines. Actes du X^e Congrès International des Linguistes, Bucarest 28 août - 2 septembre 1967, 3, Bucarest: Editions de l'Académie de la République Socialiste de Roumanie, 27-32.

Position of phrase boundaries in verse and their relation to caesurae; variation and evolution of different techniques of Catullus, Lucretius, Vergilius, Horatius, Lucanus, Iuvenalis, Ovidius, and Claudianus.

104. * HELLEGOUARC'H, J. (1976), Aspects stylistiques de l'expression de la tristesse et de la douleur dans les poèmes ovidiens de l'exil. Ovidianum. Acta conventus omnium gentium ovidianis studiis fovendis, 325-340. Cur. N. Barbu, E. Dobroiu, M. Nasta. Bucurestiis typis univ.

105. HELLEGOUARC'H, J. (1978a), La réalisation de la césure dans l'hexamètre latin. J. COLLART (ed.): Varron, Grammaire antique et stylistique latine. Paris: Les Belles Lettres, 383-395.

Place of the trihemimeral, penthemimeral, and hephthemimeral caesura of the trochaic and the bucolic diaeresis; percentages of d/s in the first four feet in Homeros and in some Latin poets (Catullus, Lucretius, Vergilius, Ovidius, Statius); in Greek hexameter the trochaic caesura dominates in the third foot, which includes most of the dactyls, whereas in Latin hexameter the trochaic caesura dominates in the first foot and the percentage of dactyls decreases from the first to the fourth foot.

106. HELLEGOUARC'H, J. (1978b), Les structures stylistiques de la poésie latine. Méthode d'analyse et application pratique. L'Information littéraire 30, 234-244.

Different percentages of dactyls in the first four feet in the hexameter of Homeros and of several Latin authors (Ennius, Catullus, Lucretius, Vergilius, Propertius, Ovidius, Iuvenalis) show that in Greek hexameters the third foot contains most of the dactyls, whereas in Latin hexameters it is the first and the fifth feet that include most of the dactyls.

107. HERDAN, G. (1954), Informationstheoretische Analyse als Werkzeug der Sprachforschung. Die Naturwissenschaften 41, 13, 293-295.

Distribution of d/s is not the outcome of chance but the outcome of word structure; this result is based on data from Drobisch and Boldrini.

108. HERDAN, G. (1962), The calculus of linguistic observations. The Hague: Mouton (= Janua linguarum 9).

Gap frequencies in Vergilius' Georgica; relative frequencies of all initial phonemes in the four books of Vergilius' Georgica; gap frequencies of selected initials (pp. 80-85).

109. HERDAN, G. (1966), The advanced theory of language as choice and chance. Berlin et al.: Springer.

Sequences of d/s in the Latin hexameter; relative frequencies of hexameters according to the number and position of d/s; distribution of combinations and permutations; comparison of patterns in Vergilius, in Vergilian poets and non-Vergilian poets; sentence length and caesurae in the early Greek hexameter (pp. 206-213).

110. HERDAN, G. (1969), The mathematical theory of verse. ZPhon 22, 225-234.

Some considerations on the factors which contribute to making the language of poetry different from that of prose. Uses material from Czech, Russian, and English.

111. HERR, M.W. (1966), The additional short syllables in Ovid. New York: Kraus Reprint Corporation (reprint) (= Lg diss. 25, 1937).

Morphological examination of Ovidius (Metamorphoseon libri 1, 2, 5, 10, 15) and Vergilius (Aeneis 1, 2, 6): frequencies of neuter nouns common in Ovidius but not common in Vergilius; percentages of frequency of all neuter nouns; ratio of plural to singular neuter nouns in case-forms; average frequency per 100 lines of final syllables in nominative, accusative, and neuter nouns with vocalic initial.

112. HEYKEN, J. (1916), Über die Stellung der Epitheta bei den römischen Elegikern. (Diss.) Kiel: Fiencke.

Position of the epitheton in relation to its noun; position of the two in verse; there are principles which indicate an intentional technique; data from Catullus, Ovidius, Tibullus, Propertius, and others.

113. HORNIG, W. (1972), Theorie einer systematischen lateinischen Metrik. (Diss.) Frankfurt/Main.

Structural laws in Horatius, Satirae (book I), Ovidius, Metamorphoseon libri, Statius, Silvae (book I), Vergilius, Aeneis, and Ennius, Annales.

114. INGALLS, W.B. (1979), Formular density in the similes of the Iliad. TAPA 109, 87-109.

An investigation of the formular density of a selection of similes and of a control passage of narrative reveals a uniform formular texture in the Ilias.

115. INGALLS, W.B. / PARKER, L.P.E. (1970), Greek metrics. Phoenix 24, 1-12.

Tables compiled from Porter's data (PORTER 1951) giving the number of word-ends in each position in the 1000 verse samples of the Ilias and Odysseia and the sequence of positions in descending order of frequency, first for the Ilias, next for the Odysseia, and finally for the combination of both.

116. JONES, F.P. (1966), A binary-octal code for analyzing hexameters. TAPA 97, 275-280.

Computational notation of the hexameter with frequencies in Homeros and Vergilius.

117. JONES, F.P. / GRAY, F.E. (1972), Hexameter patterns, statistical inference and the Homeric question: an analysis of the La Roche data. TAPA 103, 187-209.

Comparison of the Ilias as a whole with separate books of the Ilias; the Ilias is more similar to the Odysseia than to the Homeric hymns, Hesiodos, or Aratos, Kallimachos, Apollonios, and Nonnos, according to patterns of dactyls and spondees.

118. KIRK, G.S. (1966), Studies in some technical aspects of Homeric style. Yale Classical Studies 20, 73-152.

Occurrences of different metrical word-types within the verse or at verse-ends of the Ilias and the Odysseia; sentence-length in the Ilias (book 16).

119. KNIGHT, W.F.J. (1931), Homodyne in the fourth foot of the Vergilian hexameter. CQ 25, 184-194.

Data on the frequency of the fourth foot homodyne, on percentages of released movements, and alternations in Vergilius, Tibullus, Catullus, Lucretius, Ovidius, Lucanus, and Valerius Flaccus.

120. KNIGHT, W.F.J. (1950), Accentual symmetry in Vergil. Oxford: Blackwell (2nd edition).

Tables of fourth foot texture in Vergilius; data on Bucolica 1-9, Georgica 1-4, Aeneis 1-12, compared with data on Lucretius' De rerum natura 1, Catullus' Carmina 64, Tibullus' Elegiarum libri 4,1, Ovidius' Metamorphoseon libri 1, Lucanus' Pharsalia 1, Valerius Flaccus' Argonautica 1, and with data on Culex, Ciris, Moretum, Dirae, Lydia, Est et non, Vir bonus, Aetna.

121. KOESTERS, H. (1893), Quaestiones metricae et prosodicae ad Valerium Flaccum pertinentes. (Comm. philol.) Münster.

Numerical information about elisions (absolute numbers and percentages) in single books and single verses of the Argonautica (books 1-8).

122. KOLLMANN, E.D. (1968), Remarks on the structure of the Latin hexameter. Glotta 46, 293-316.

The fact that words which retain their prose accent in verse have been found to be the key words of the verse, especially those bearing ictus and accent on a long syllable, seems to prove the basic assumption that both ictus and accent are of dynamic nature; data on Vergilius' Aeneis 1, 1-11 and 2, 40-49.

123. KOLLMANN, E.D. (1974), A new method of analyzing Latin hexameter. Scripta classica israelica. Yearbook of the Israel society for the promotion of classical studies 1. Jerusalem: Academic press, 64-72.

Result of his "qualified metrical analysis" (i.e. distinction between six various "sound-levels"): some 80% of the words in the Latin hexameter retain their prose-accentuation (including ictus-bearing monosyllables), only about 20% have a voluntary clash; this seems to show that the Latin hexameter is based mainly on coincidence of ictus and accent; examination of a random sample of about 600 lines by various Latin poets.

124. KOLLMANN, E.D. (1975a), Konsonantenfolge im Lateinischen und ihre syntaktische Bedeutung. Glotta 53, 101-121.

Analysis of 7083 word breaks in various authors; data on consonant clusters in all positions; there is a connection between syntax and type of consonant sequence, many of which are rare.

125. KOLLMANN, E.D. (1975b), The infinitive in Latin hexameter poetry. Glotta 53, 281-291.

Manually collected data on the prosody and the placement of infinitives; hypothesis that the structure of the infinitive form influences its location in the verse and indirectly its syntactical function.

126. KOLLMANN, E.D. (1976), Observations on the occurrence of the parallel forms a = ab and e = ex in Latin literature. RhM 119,1, 14-29.

Comparative frequency of occurrence (in prose and poetry) of a/ab and e/ex in Ennius, Plautus, Terentius, Lucretius, Lucanus, Catullus, Tibullus, Propertius, Vergilius, Horatius, Ovidius, Statius, and others.

127. KOLLMANN, E.D. (1977), Some more information on a and ab. RhM 120, 346-364.

Tables of the relative frequencies of a/ab in Ovidius' elegies.

128. KUSCH, A. (1915), De saturae Romanae hexametro questiones historicae. (diss. in.) Greifswald: Typis Roberti Noske Bornensis.

Distribution of elisions in single feet (breves, mediae, longae); data on the Saturae (book 1 and 2) of Horatius compared with data on other Latin authors.

129. LAPŠINA, N.V. / ROMANOVIČ, J.K. / JARXO, B.J. (1934), Metričeskij spravočnik k stixotvorenijam A.S. Puškina. Moskva, Leningrad: Akademia.

Percentage of verses with replacement of dactyls by trochees in Puškin's works according to feet and creative periods. The percentage of verses without replacement is given as well.

130. LA ROCHE, J. (1898a), Zahlen-Verhältnis im homerischen Vers. WS 20, 1-69.

Frequency distribution of d/s in the verses of Homeros (Ilias, Odyseia).

131. LA ROCHE, J. (1898b), Untersuchungen über den Vers bei Hesiod und in den homerischen Hymnen. WS 20, 70-90.

Comparison of d/s distribution in Homeros and in Hesiodos.

132. LA ROCHE, J. (1899), Der Hexameter bei Apollonius, Aratos und Kallimachos. WS 21, 161-197.

Frequency and placement of hephthemimeral, penthemimeral, and main caesura, numbers of spondees in single feet.

133. LA ROCHE, J. (1900a), Zur Prosodie und Metrik der späteren Epiker. WS 22, 35-55.

Number of verses containing 3 spondees; data from Quintius, Kolluthos, Nikandros, Oppianos, Manethos, and Tryphiodoros.

134. LA ROCHE, J. (1900b), Zur Verstechnik des Nonnos. WS 22, 194-221.

Number of verses according to their relation of d/s; data from the first five books of the Dionysiaka.

135. LA ROCHE, J. (1901), Der Hexameter bei Vergil. WS 23, 121-142.

Stylistic comparison of Vergilius and Homeros concerning the number and the position of d/s; statistical analysis of the hephthemimeral caesura.

136. LECROMPE, R. (1970), Virgile, Bucoliques. Index verborum. Relevés statistiques. Hildesheim, New York: Olms (= Alpha Omega 24. Lexika-Indizes-Konkordanzen zur Klassischen Philologie).

Distribution of the vocabulary according to grammatical categories; forms of the verb: frequency distribution according to mode, tense, and voice.

137. LEGRAND, P.E. (1898), Etude sur Théocrite. Paris: Bibliothèque des Ecoles Françaises d'Athènes et de Rome.

Analysis of the distribution of the elision between the six feet of the hexameter; numbers of the semiquinaria in the verse of Theokritos; data from the Pharmakeutria.

138. LEVISON, M./ MORTON, A.Q./ WINSPEAR, A.D. (1968), The Seventh Letter of Plato. Mind 77, 309-325.

Some comparative data on sentence length distribution,as wellas on the occurrence of καί/δέ, and on the number of works and sentences in some works of Plato to determine the authorship of the Seventh Letter.

139. LIÉNARD, E. (1978), Répertoires prosodiques et métriques: Lucrèce, De rerum natura, L.III, Valerius Flaccus, Argonautica, L.VIII, Germanicus, Aratea. Bruxelles: Editions de l'Université de Bruxelles.

Three prosodic and metrical analyses founded on the same basic principles as those that have led L. Nougaret with his Analyse verbale comparée du De Signis et des Bucoliques (NOUGARET 1966).

140. LONGPRÉ, A. (1972), Traitement de l'élision chez le poète Cyprianus Gallus. Phoenix 26, 63-67.

Manual analysis according to the study of SOUBIRAN (1966); position of elision; confirmation of Soubiran's assumption that grammatically important elements are not affected by elision as frequently as Soubiran claims.

141. LONGPRÉ, A. (1975), Aspects de métrique et de prosodie chez Juvencus. Phoenix 29, 128-138.

Percentages of d/s in the hexameter of Iuvencus, Vergilius, Ovidius, Lucanus, Statius, Cyprianus, and Paulinus de Pella.

142. * LUCOT, R. (1955), Un type d'hexamètre latin d'Ennius à Virgile. Pallas 3, 29-39.

143. * LUCOT, R. (1965), Ponctuation bucolique, accent et émotion dans l'Enéide. REL 43, 261-274.

Numerical information about Vergilius' Aeneis.

144. * LUCOT, R. (1969), Sur l'accent de mot dans l'hexamètre latin. Pallas 16, 79-106.

145. LUDWICH, A. (1866), De hexametris poetarum Graecorum spondaicus. Halis: Typis Orphanotrophei.

Distribution of d/s in Greek authors, especially in the Ilias and the Odysseia.

146. LUDWICH, A. (1971), Aristarchs homerische Textkritik. Nach den Fragmenten des Didymos. Band II. Hildesheim/New York: Georg Olms, reprografischer Nachruck der Ausgabe Leipzig 1885.

Table of the numerical relation of d/s in the first five feet of Greek hexameter; number of verses with given number of spondees contained; total frequency of the different forms of verse in Homeros.

147. LÜDTKE, H. (1968), Metrik und Informationstheorie. J. Levý, K. Palas (eds.): Teorie verše II. Sborník druhé Brněnské versologické konference (18. - 20. řijna 1966). Brno: Universita J.E. Purkyně, 61-66.

Description of metrical redundancy-calculation and its application to the beginning of Vergilius' Aeneis.

148. LÜDTKE, H. (1969), Der Vergleich metrischer Schemata hinsichtlich ihrer Redundanz. H. KREUZER/ R. GUNZENHÄUSER (eds.) Mathematik und Dichtung. München: Nymphenburger Verlagshandlung (3rd edition), 235-237.

Redundancy in hexameter verse; redundancy amount per syllable in the dactylic hexameter.

149. MAAS, H.D. (1971), Einige statistische Untersuchungen zum Werk Georg Trakls. LiLi 4, 43-50.

Statistical evaluation of phoneme frequencies in the poetic works of Trakl.

150. MARTIN, W. (1973), A note on Herdan's theory on the recurrence of initial phonemes. RELO, fasc. 2, 1-17.

Probabilities of occurrence of initial phonemes in Vergilius' Georgica.

151. MERRIL, W.A. (1922), The Lucretian hexameter. University of California publications in classical philology 5, 253-296. Berkeley: Berkeley University press.

Distribution of d/s, frequencies of elisions and caesurae in Lucretius.

152. MERRIL, W.A. (1924), The metrical technique of Lucretius and Cicero. UCPL 7, 293-306.

Comparison between the hexameter of Lucretius and that of Cicero; data on the penthemimeral and hephthemimeral caesura, on the diaeresis; analysis of the distribution of d/s; examination of the end of verse; data on elision and synaloepha.

153. MICHAELSON, S./MORTON, A.Q./ WAKE, W.C. (1978), Sentence length distributions in Greek hexameter verse and Homer. Association for literary and linguistic computing bulletin vol. 6, no. 3,254-267.

An exhaustive study of the constraints imposed on sentence length by hexameter verse forms; includes data from Homeros, Hesiodos, Apollonios Rhodios, and Aratos. Many sentences cover exactly one or two lines, but the distributions do not distinguish between authors; proportions of sentences of different places in the line tend to distinguish primitive and literary compositions and also divide the Odysseia (24) into three parts.

154. MICHEL, J.-H. (1970), Une hypothèse de travail sur les rapports entre l'ictus et l'accent dans l'hexamètre latin. RELO, fasc. 3, 1-17.

Examination of the relationship between word accent and verse ictus for part of Vergilius' Bucolica I, assuming a pitch accent.

155. MIKS, A. (1937), Heksameeter ja eleegiline distihhon eesti teoorias ja värsis. Eesti Kirjandus 2, 109-118 and 3, 153-165.

Distribution of d/s patterns in 100 verses of the Aeneis and of the Ilias; comparison with data on the hexameter of the Finnish translation of the Ilias by O. Manninen, Estonian translations of the Ilias by J. Jõgever, V. Ridala, J. Lõo, and J. Bergman and with M.J. Eisen's narrative poem "Kõu and Pikker". Numerical information about the history of Estonian hexameter. Frequencies of quantity mistakes (short for long and long for short) in the elegiac verses of G. Suits.

156. MINYARD, J.D. (1978), Mode and value in the De Rerum Natura. A study in Lucretius' metrical language. Wiesbaden: Steiner (= Hermes, Einzelschriften 39).

Data concerning verses containing metrically stable expressions; statistics on verses with which sentences end; tables of beginning positions for non-initial phrase formulae and of ending positions for the phrase formulae; numbers of word repetitions and word combinations.

157. MOERNER, F. (1890), De P. Papinii Statii Thebaide quaestiones criticae, grammaticae, metricae. (diss. in.) Königsberg.

Distribution of d/s (absolute numbers and percentages); comparison with data of Drobisch (DROBISCH 1866) which differ from those of Moerner.

158. MORTON, A.Q. (1979), Literary detection, chapter on the Homer problem. London: Bowler. New York: Scribners.

A simple account of the Michaelson, Morton and Wake study (MICHAELSON, MORTON, WAKE 1978), with some notes on how to take advantage of constraints laid upon a writer by literary form. Such constraints invariably produce other features by which identification of the writer can be achieved.

159. MORTON, A.Q./ MICHAELSON, S. (1977), Watch this space. A cautionary tale for classical scholars. RELO, fasc. 1, 51-61.

It is not possible to draw valid conclusions about the authenticity of a text by considering the frequency of a term only and neglecting at the same time strict mathematical analysis of the text.

160. NEUBOURG, L. de (1977), Mots longs après les diérèses médianes de l'hexamètre latin. Pallas 24, 45-79.

Data on frequencies of long words after the middle diaeresis, as well as on polysyllabic word endings in the second and the third foot, and on the phonetic treatment of the diaeresis in Calpurnius, Martialis, Ovidius, Lucanus, and Iuvenalis.

161. NEUBOURG, L. de (1978a), Mots longs après les diérèses médianes dans l'hexamètre grec. L'Antiquité classique 47, 403-437.

Numerical data on metrical types of long words after the third foot, on the syllabic connection of long words, on the syllabic connection of dactylic and spondaic words, and on the distinction of elisions in different Greek authors.

162. NEUBOURG, L. de (1978b), Taaldwang en metrische dwang bij de plaatsing van adjectief en substantief in de Latijnse dactylische hexameter. Een grammaticaal en metrisch onderzoek, toegepast op Vergilius' Georgica. Glotta 56, 102-122.

Abstract of the author's (unpublished?) dissertation with reports on the influence that word ending, word stress, word length, the phonetic structure of the word, and the grammatical class have on the poetic word order; assumption that the prosodic structure is characteristic of metrics.

163. NILSSON, N.-O. (1952), Metrische Stildifferenzen in den Satiren des Horaz. Uppsala: Almqvist & Wiksells (= Studia Latina Holmiensia 1).

Comprehensive analysis of different sorts of elisions, of caesurae, of punctuation, and of enjambement in the first two Satirae.

164. NILSSON, N.-O. (1960), Verschiedenheiten im Gebrauch der Elision in Vergils Eklogen. Eranos 58, 80-91.

The fact that in the Bucolica there are less numerous elisions than in Vergilius' larger works is an illuminating example of differences in using elisions dependent on the literary genre; analysis of the homogeneity of elisions within the Bucolica; frequencies of elisions (Bucolica 1-10); percentages of verses with several elisions or with elided longae.

165. NOUGARET, L. (1946), Les fins d'hexamètre et l'accent. REL 24, 261-271.

Some data on the verse end in Lucretius, Vergilius, and Ovidius.

166. NOUGARET, L. (1966), Prosodie, métrique et vocabulaire. Analyse verbale comparée du "De Signis" et des "Bucoliques". Paris: Les Belles Lettres.

Metrical arrangements of words which are interchangeable with regard to their number of syllables.

167. OLLFORS, A. (1967), Studien zum Aufbau des Hexameters Lucans. (= Acta Regia Societatis Scientiarum et Litterarum Gothoburgensis. Humaniora 1).

Analysis of the position of different metrical word types in the hexameter of Lucanus; table of all words with initial a and their position in the verse; examination of words with fixed places in the hexameter; study of the role these fixed words in Lucanus play in the works of authors from Ennius to Iuvenalis.

168. O'NEILL, E.G. (jr.) (1940), Word-accents and final syllables in Latin verse. TAPA 71, 335-359.

Some data on the occurrence of d/s in the first foot in the first 1000 verses of the Aeneis of Vergilius and Ovidius' Metamorphoseon libri.

169. O'NEILL, E.G. (jr.) (1942), The localization of metrical word types in the Greek hexameter: Homer, Hesiod, and the Alexandrians. Yale Classical Studies 8, 103-178.

Analysis of 7152 hexameters from Homeros, Hesiodos, Aratos, Kallimachos, Apollonios, and Theokritos; patterns of long and short syllables within a single word are tabulated according to their position in the line; quantitative analysis of positions filled up by different metrical types in the verse; frequencies of d/s; frequencies of word beginnings and word ends.

170. OTT, W. (1966) Metrical analysis of Latin hexameter by computer, I. RELO, fasc. 4, 7-24.

Description of a computer analysis program.

171. OTT, W. (1967a), Metrical analysis of Latin hexameter by computer, II. RELO, fasc. 1, 39-64.

Demonstration of a complete computational analysis of a poem choosing as an example the short poem Est et non (Appendix Vergiliana).

172. OTT, W. (1967b), Metrical analysis of Latin hexameter by computer. Corrigenda. RELO, fasc. 2, 48-49.

173. OTT, W. (1970), Metrische Analysen zur Ars Poetica des Horaz. Göppingen: Kümmerle (= Göppinger Akademische Beiträge Nr. 6).

Tables showing the position of elision in the verse; coincidence of ictus and word accent; verse endings, and metrical word types.

174. OTT, W. (1973a), Metrische Analysen zu Vergil, Aeneis Buch VI. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 1).

Information on metrical characteristics, word length, frequencies of word stress in single feet, different types of elision, word endings in single feet, frequencies of colon boundaries in single feet, rhythmical types, spondaic words in the first foot, types of verse endings and types of word endings respectively, metrical types; metrical and alphabetical word index; the punched cards included permit manual determination of lines with words ending in particular positions; exhaustive analysis.

175. OTT, W. (1973b), Metrische Analysen zu Vergil, Aeneis Buch I. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 2).

Analysis as in OTT 1973a.

176. OTT, W. (1973c), Metrische Analysen zu Vergil, Aeneis Buch XII. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 3).

Analysis as in OTT 1973a.

177. OTT, W. (1973d), Metrische Analysen zu Catull, Carmen 64. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 4).

Analysis as in OTT 1973a.

178. OTT, W. (1973e), Metrische Analysen zu Statius, Thebais Buch I. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 5).

Analysis as in OTT 1973a.

179. OTT, W. (1974a), Metrische Analysen zu Lukrez, De rerum natura, Buch I. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 6).

Analysis as in OTT 1973a.

180. OTT, W. (1974b), Metrische Analysen zu Ovid, Metamorphoseon, Buch I. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 7).

Analysis as in OTT 1973a.

181. OTT, W. (1974c), Rückläufiger Wortindex zu Vergil Bucolica, Georgica, Aeneis. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 8).

Information on the position of words in the line and on the quantity of their syllables to demonstrate metrical parallels.

182. OTT, W. (1978a), Metrische Analysen zu Vergil, Bucolica. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 9).

Analysis as in OTT 1973a.

183. OTT, W. (1978b), Metrische Analysen zu Vergil, Georgica I: Analysen und Übersichten. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 10).

Analysis as in OTT 1973a.

184. OTT, W. (1978c), Metrische Analysen zu Vergil, Georgica II: Indizes. Tübingen: Niemeyer (= Materialien zu Metrik und Stilistik 11).

Analysis as in OTT 1973a.

185. PACKARD, D.W. (1974), Sound-patterns in Homer. TAPA 104, 239-260.

Tables of lines with unusually high frequencies of given sounds, unusual combinations of sounds at word boundaries; presentation of a formula for recognizing smooth and harsh verses by Dionysius' rules.

186. PACKARD, D.W. (1976), Metrical and grammatical patterns in Greek hexameter. A. JONES and R.F. CHURCHHOUSE (eds.). The computer in literary and linguistic studies. Pro-

ceedings of the third international symposium. University of Wales, Cardiff, 85-91.

The comparison of structural formulae (Odysseia 1, Quintus Smyrnaeus' Posthomerica 12), which are defined as repeated grammatical patterns in fixed metrical positions, permits the following conclusion: the frequency of structural formulae in Homeros does not clearly exceed that found in any other literary style.

187. PARRY, M. (1929), The distinctive character of enjambement in Homeric verse. TAPA 60, 200-220. Reprinted in A. PARRY (ed.). The making of Homeric verse, Oxford 1971, 251-265.

The different sorts of enjambement in Homeros are compared to those found in Vergilius and Apollonius. In opposition to Vergilius and Apollonius Homeros ends his thoughts most frequently within the line and he characterizes more intensely the rhythm of the hexameter. Analysis of six passages of one hundred lines each from the Ilias, the Odysseia, the Aeneis, and the Argonautica; tables on verses which are followed by unperiodic enjambement and on verses after which enjambement is necessary.

188. PARRY, M. (1971), The distinctive character of enjambement in Homeric verse. A. PARRY (ed.): The making of Homeric verse. The collected papers of Milman Parry. Oxford: Clarendon press, 251-265.

Analysis as in PARRY 1929.

189. PAULSON, J. (1897), Lukrezstudien I. Die äussere Form des Lukretianischen Hexameters. Gothenburg: Wettergren & Kerber (= Göteborgs Högskolas Årsskrift 15).

Absolute numbers and percentages of d/s and numerical information about verses with one, two, three or four spondees in Lucretius (Elegiae 1-6); comparison of patterns in Lucretius and Vergilius (Aeneis).

190. PERRET, J. (1954), Mots et fins de mots trochaïques dans l'hexamètre latin. REL 32, 183-199.

Analysis of restrictions of the word structure in the trochaic verse of Lucretius, Cicero, Catullus, Vergilius, Tibullus, Propertius, Manilius, Lucanus, Valerius Flaccus, Silius, Claudius, Prudentius, Horatius, and Ovidius.

191. PERRET, J. (1955), Le partage du demi-pied dans les anapestiques et dans l'hexamètre. REL 33, 352-366.

Data from Aristophanes, Plautus, Ennius, Pacuvius, Accius, Seneca, Lucretius, Catullus, Ovidius, Cicero, Horatius, Lucanus, Valerius Flaccus, Homeros, and Apollonios Rhodios.

192. PERRET, J. (1956), Ponctuation bucolique et structure du IVe pied. REL 34, 146-158.

Data from Homeros (Ilias), Apollonios Rhodios, Aratos, Vergilius (Georgica, Aeneis), Ovidius (Metamorphoseon libri), Valerius Flaccus, and Iuvenalis.

193. PERRET, J. (1957), Prosodie et métrique chez Commodien. Pallas 5, 27-42.

Data on long syllables with the function of short syllables and on short syllables with the function of long syllables (1060 verses of Commodianus' Carmen apologeticum).

194. PIPPING, H. (1938), Zur homerischen Metrik; eine statistische Untersuchung. Societas Scientiarum Fennica: Commentationes Humanarum Litterarum 9, fasc. 6, 1-11.

Analysis of the syllabification in the Ilias and the Odysseia of Homeros.

195. PIPPING, H. (1943), Ueber die Bedeutung der Wortgrenzen in der römischen und griechischen Metrik. Societas Scientiarum Fennica: Commentationes Humanarum Litterarum, 13, fasc. 5.

The frequency of free feet in the first and the third position is higher in Horatius (Satirae, Epistulae) than in Vergilius (Aeneis I-V, VI 1-254); comparison with the Odysseia I of Homeros.

196. PLATNAUER, M. (1948), Elision of atque in Roman poetry. CQ 42, 91-93.

Percentages of elided and non-elided ille in Vergilius, Catullus, Ovidius, Tibullus, and Propertius; data on the elision of atque in Lucretius, Ovidius, Lucanus, Statius, Valerius Flaccus, Silius Italicus, Horatius, Iuvenalis, Persius, and Martialis.

197. PLATNAUER, M. (1951), Latin elegiac verse. Cambridge: Cambridge University press.

Frequency and distribution of d/s in Tibullus, Propertius, and Ovidius; some data on elision of short open vowels, of -m, of long open vowels, of diphthongs, and of monosyllables.

198. POE, J.P. (1974), Caesurae in the Latin hexameter line of Latin elegiac verse. Wiesbaden: Steiner (= Hermes, Einzelschriften 29).

Examination of hexameter lines in all of the elegiac verses of Propertius, Ovidius, and the Corpus Tibullianum; position of word ends, clause ends, and noun-adjective phrases within the verse.

199. PORTER, H.N. (1951), The early Greek hexameter. Yale Classical Studies 12, 1-63.

Description of such formular elements as are common to all early Greek hexameter verse: 1. classification according to the cola of which lines are composed, 2. composition of each colon in terms of the metrical word-types of which it is composed, 3. presentation of figures for the metrical word-types which end in each of the possible positions for word-ends in the line; data from Homeros, Hesiodos, and Kallimachos.

200. RADFORD, R.S. (1923), Tibullus and Ovid. The authorship of the Sulpicia and Cornutus Elegies in the Tibullan Corpus. AJPh 44, 1-26, 230-259, and 293-318.

Comparison of d/s combinations occurring in the verses of Ovidius and Tibullus; dssd is the most preferred by Ovidius which distinguishes him from all other elegiac poets; dssd is relatively infrequent in Tibullus; the great Tibullan schema dsds is relatively infrequent in the youthful Ovidius.

201. * RASI, P. (1897), In Claudii Rutilii Namatiani"De redictu suo"libros adnotationes metricae. Rivista di filologia e di istruzione classica 25, 169-214.

202. REUTER, E. (1909), De Avieni hexametrorum re metrica. (diss. in.) Bonn.

Data on the proportion of spondees, monosyllables, tetrasyllables, pentasyllables in several Latin authors; data on the caesura semiquinaria, on proportions of d/s, and on syllables ending in vowel or in -m in Avienus' Carmina 2 and 3.

203. RICHMOND, J.A. (1965), A note on the elision of final e in certain particles used by Latin poets. Glotta 43, 78-103.

The purpose of this article is to demonstrate statistically that with a number of words a tendency to use unelided forms less frequently than might be expected is observable in many Latin poets.

204. RICHMOND, J.A. (1968a), Metre and prosody in the "Halieutica" ascribed to Ovide. Hermes 96, 341-355.

Elisions, caesurae, and distribution of dactyls in the Halieutica and the Metamorphoseon libri of Ovidius.

205. RICHMOND, J.A. (1968b), -que que- in classical Latin poets. Philologus 112, 135-139.

Numbers of words with initial que- before which -que could have stood or stands; data from Catullus, Horatius, Tibullus, Phaedrus, Seneca, Persius, Statius, Iuvenalis, Lucretius, Vergilius, Propertius, Ovidius, Lucanus, Valerius Flaccus, Silius Italicus, and Martialis.

206. SCHER, L.O. (1972), The structure of Lucan's hexameter. (diss.) Stanford.

Positions of different caesurae in Lucanus' Bellum civile 1.

207. SCHMIEL, R. (1981), Rhythm and accent: Texture in Greek epic poetry. R. GROTJAHN (ed.): Hexameter studies. Bochum: Brockmeyer, 1-32 (= Quantitative linguistics 11).

Study on the relation between ictus and accent in Greek epic poetry with data primarily from Homeros; conclusion that vivid lines tend to be homodyne whereas calm lines tend to be heterodyne.

208. SCHULTE, G. (1909), De Ovidiano synaloepharum usu. (diss. in.) Münster.

Absolute numbers and percentages of synaloepha in Ovidius (Heroides, Amores, Remedia amoris, Metamorphoseon libri, Fasti, Tristium libri 5, Ibis, Epistulae ex Ponto and others) compared with data from Tibullus (Elegiarum libri 4) and Propertius (Elegiarum libri 4).

209. SHEWAN, A. (1925), Alliteration and assonance in Homer. CPh 20, 193-210.

Quantitative analysis demonstrating that Homeros used alliteration.

210. SIEDOW, A. (1911), De elisionis aphaeresis hiatus usu in hexametris Latinis ab Ennii usque ad Ovidii tempora. (diss. in.) Greifswald.

Comprehensive statistical tables on elision, aphaeresis, and hiatus; data from Ennius, Lucilius, Cicero, Lucretius, Catullus, Vergilius, Horatius, Tibullus, Propertius, and Ovidius.

211. SKUTSCH, F. (1901), Aus Vergils Frühzeit. Leipzig: Teubner.

Frequencies of synalophae after certain feet in Lucretius, Catullus, Vergilius, and Ovidius (with no indication of the number of the verses examined).

212. SOUBIRAN, J. (1955), L'hexamètre de Cicéron. Pallas 3, 41-59.

Support for hypothesis on the development of the Ciceronian technique based on an examination of the hexameter ends in Cicero as compared with Ennius, Lucretius, Vergilius, Ovidius, Lucanus, and Catullus.

213. SOUBIRAN, J. (1956), L'élision et la structure du vocabulaire latin. REL 34, 39-40.

Comparison of probabilities and observed frequencies of elision in Vergilius and Lucanus.

214. SOUBIRAN, J. (1965), Prosodie et métrique des "Bella Parisiacae Vrbis" d'Abbon. Journal des savants, 204-331.

Stylistical comparison between Abbon and Latin hexameter poets concerning the quantity of vowels and syllables, elisions, and the frequencies of d/s in single feet.

215. SOUBIRAN, J. (1966), L'élision dans la poésie latine. Paris: Klincksieck.

Frequencies of elision according to feet in Lucretius, Vergilius, Horatius, Ovidius, and Lucanus; types of words in contact with elision in Ennius, Vergilius, and Lucanus; numbers of elided ă, ĕ, ā, ō, ī, in the Aeneis of Vergilius; numbers of elided iambus in Ennius, Cicero, Lucretius, Vergilius, Horatius, Ovidius, and Lucanus.

216. SOUBIRAN, J. (1968), L'hexamètre latin. Problèmes de structure et de diction. REL 46, 410-424.

Some absolute numbers on the sequence of consonant-vowel and vowel-consonant in the bucolic punctuation in Vergilius, Ovidius, and Lucanus.

217. SOUBIRAN, J. (1978), Monosyllabes introducteurs devant la césure: Ennius, Plaute et leurs modèles grecs. J. COLLART (ed.): Varron, grammaire antique et stylistique latine. Paris: Les Belles Letres, 321-336.

Frequencies of elidable final short and long syllables, and -m in Plautus' Aulularia (2371 words) and Mercator (2895 words).

218. STEELE, R.B. (1911), The endings -ere and -erunt in dactylic hexameter. AJPh 32, 328-332.

Raw data on the occurrence of -ere/-erunt in the six feet.

219. STEELE, R.B. (1926), Variation in the Latin dactylic hexameter. PhQ 5, 212-225.

Enumeration of all possible combinations of d/s in Ennius, Catullus, Cicero, Horatius, Persius, Iuvenalis, Silius Italicus, Lucretius, Vergilius, Manilius, Lucanus, Statius, Ovidius, and Valerius Flaccus.

220. STEELE, R.B. (1929), Elision in Latin dactylic hexameter. PhQ 8, 43-60.

Tables of elided est, -que, -m, longum, and breve; list of occurrences of per in Horatius and Vergilius.

221. STIFLER, Th. (1924), Das Wernickesche Gesetz und die bukolische Dihärese. Philologus 79, 323-354.

Result of his analysis of the Aeneis and the Odysseia: Wernicke's law has no validity for Homeros, because Homeros did not intentionally avoid long syllables in the fourth thesis, emerging in word juncture position.

222. STURTEVANT, E.H. (1919), The coincidence of accent and ictus in the Roman dactylic poets. CPh 14, 373-385.

A comparison of the relationship between accent and ictus in prose writers and dactylic poets: Ennius, Catullus, Lucretius, and Vergilius.

223. STURTEVANT, E.H. (1921a), On the frequency of short words in verse. The Classical Weekly 15, 73-76.

Some of the Roman poets use a smaller proportion of monosyllables than occur in Latin prose, while English verse usually has more monosyllables than prose. Tables provide information about the number of syllables and the frequency of mono- and disyllables in Greek, Latin and English authors.

224. STURTEVANT, E.H. (1921b), Word-ends and pauses in the hexameter. AJPh 42, 289-308.

Position of word-ends in the first 500 verses of Homeros' Ilias and Vergilius' Aeneis; lists of:
- words with which a verse may end and their proportion in verse as a whole
- proportions of Homeros' words of each of the rhythmic types possible before each of the possible types of verse-finals
- points per 100 lines in Homeros, Hesiodos, Theokritos, Vergilius, Lucretius, Horatius, and Ovidius
- proportion of all points in each position.

225. STURTEVANT, E.H. (1923), Harmony and clash of accent and ictus in the Latin hexameter. TAPA 54, 51-73.

Percentages of words bearing ictus which show harmony in Ennius, Lucilius, Catullus, Lucretius, Cicero, Vergilius, Horatius, Ovidius, Lucanus, Persius, Iuvenalis, Statius, and Silius Italicus in relation to ictus syllables of known character which show harmony or sequences of harmony, clash of accent and ictus in 2nd, 3rd, and 4th feet.

226. STURTEVANT, E.H. (1924a), Accent and ictus in the Latin elegiac distich. TAPA 55, 73-89.

Comparison of hexameter and pentameter with a table on the hexameter from STURTEVANT 1923.

227. STURTEVANT, E.H. (1924b), The doctrine of caesura, a philological ghost. AJPh 45, 329-350.

Number of occurrences per 100 lines of word ends coinciding with the ends of each of the first five feet; data from 800 lines of the Ilias and the Aeneis, 500 lines of Theokritos, and 300 lines of Apollonios, Catullus, and Lucretius.

228. STURTEVANT, E.H. (1939), Horace, Carm. 3.30. 10-14, and the Sapphic Stanza. TAPA 70, 295-302.

Tables which show the frequency of occurrence of words bearing the accent on position 1, 4, 6, 10 in the line in Catullus, Horatius, Seneca, Statius, and Ausonius.

229. STURTEVANT, E.H./ KENT, R.G. (1915), Elision and hiatus in Latin prose and verse. TAPA 46, 129-155.

Percentages and absolute numbers of elisions in Ennius, Lucretius, Catullus, Vergilius, Horatius, Tibullus, Propertius, Ovidius, Manilius, Persius, Lucanus, Calpurnius Sic., Lucilius Iunior, Valerius Flacc., Martialis, Statius, Silius Italicus, Iuvenalis, Nemesianus, and Claudianus.

230. THRAEDE, K. (1978), Der Hexameter in Rom. Verstheorie und Statistik. München: C.H. Beck (= Zetemata 71).

Distribution of d/s; word junctures; word form and metre; caesurae and metrical cola; synaloepha; choice of words; syntax; statistical methods: average, change coefficient, scattering coefficient, correlation; data from Lucretius,Vergilius, Ovidius, Lucanus, Silius, and others. Contains an annotated bibliography on the hexameter with numerous incorrect or incomplete details.

231. TORDEUR, P. (1970), Etude des élisions dans la Mosella d'Ausone. Latomus 29, 966-987.

Elisions in Ausonius' Mosella compared with elisions in Lucretius (De rerum natura), Vergilius (Aeneis 1), Horatius (Epistulae 1, 1-2), Ovidius (Metamorphoseon libri), and Lucanus (Bellum civile 3).

232. TORDEUR, P. (1971), Métrique et statistique. RELO, fasc. 4, 1-7.

First results of the statistical analyses of Latin poetry carried out at the University of Brussels; points out that the distinction between different literary genres only - and not the distinction between different periods - corresponds to differences in the use of word types.

233. TORDEUR, P. (1972a), Elisions de mots iambiques et anapestiques dans l'hexamètre latin. Latomus 31, 105-129.

Analysis of elidable non-elided and elided iambic words in 16 authors, especially in Ausonius, Claudianus, Mamertus, and Prudentius.

234. TORDEUR, P. (1972b), Essai d'analyse statistique de la métrique de Prudence. RELO, fasc. 2, 19-37.

Analysis of different word types ranked according to their length in Prudentius' Psychomachia.

235. TORDEUR, P. (1974), Elision de mots pyrrhiques et tribaques dans l'hexamètre latin. Latomus 33, 353-371.

Presentation of manually gathered statistics on the position of elision in verse from Lucretius to Sidonius.

236. TORDEUR, P. (1981), Etude statistique sur l'hexamètre d'Ausone. R. GROTJAHN (ed.): Hexameter studies. Bochum: Brockmeyer, 75-96 (= Quantitative linguistics 11).

Frequency analysis of the different metrical word types in several poems of Ausonius with the help of χ^2 tests.

237. VERRES, P. (1888), De Tib. Silii Italici Punicis et Italici Iliade Latina quaestiones grammaticae et metricae. (Comm. philol.) Münster.

Absolute numbers and percentages of the proportions of d/s in both works.

238. VROOM, H. (1927), Armaque et similia in versu hexametro Latino. Mn 44, 420-424.

The type armaque occurs more frequently - in contradistinction to the type gaudia - in the first foot than in the fifth; in the first foot the type armaque is rather unusual; this fact shows no relationship to word accent but is due to the effort to avoid the diaeresis.

239. WACHT, M. (1977), Ein teilautomatisches Verfahren zur Herstellung lexikalischer Indizes zur lateinischen Hexameterpoesie. Sprache und Datenverarbeitung 1, fasc. 1, 54-67.

Description of a system for morphological analysis developed at the University of Regensburg and applied particularly to Latin hexameters.

240. WACHT, M. 1980a), Lemmatisierter Index zu Claudian mit statistischen Anhängen zu Sprache und Metrik. Nürnberg: MCS Verlag (= Regensburger Microfiche Materialien. Texte und Arbeitsmittel zur Sprach- und Literaturwissenschaft).

Information on the number and position of lexical units; frequency classification of lexemes with percentages of rare and frequent types of words; frequencies of word classes and of inflexion.

241. WACHT, M. (1980b), Lemmatisierter Index zu Lukan mit statistischen Anhängen zu Sprache und Metrik. Nürnberg: MCS Verlag (= Regensburger Microfiche Materialien. Texte und Arbeitsmittel zur Sprach- und Literaturwissenschaft).

Analysis as in WACHT 1980a.

242. WACHT, M. (1980c), Lemmatisierter Index zu Vergil mit statistischen Anhängen zu Sprache und Metrik. Nürnberg: MCS Verlag (= Regensburger Microfiche Materialien. Texte und Arbeitsmittel zur Sprach- und Literaturwissenschaft).

Analysis as in WACHT 1980a.

243. WAITE, St.V.F. (1970), Metrical indices to Statius' Achilleid. Hephaistios 1,1, 28-70.

Tables of lines sorted by scansion; index of metrical word types sorted by shape, and lists of metrical word types sorted by position and frequency.

244. WEBER, C.W. (1975), Tmesis, sandhi, and the two caesurae of the hexameter in Vergil's "Aeneid". (diss.) University of California, Berkeley (= Language and Literature Classical).

Examines the number of lines of the Aeneis in which are found the two caesurae characteristic of Homeros, i.e. the "penthemimeral" after the third arsis, and the "Caesura tertia trochaica" after the third trochee. Raises the question whether these lines are numerous enough to justify a hypothesis on the Vergilian caesura that is as simple as that on the Homeric caesura. Sample of 1.500 lines composed of Aeneis 1 and the first 750 lines of Aeneis 12 (four incomplete lines and 12, 612f. excepted).

245. WEBSTER, T.B.L. (1941), A study of Greek sentence construction. AJPh 62, 385-415.

The purpose of this study is to attempt to determine how sentence construction changes during the period from Homeros to Deinarchos and what the individual characteristics of the different authors' sentences are.

246. WHITEHEAD, Ph.B. (1930), A new method of investigating the caesura in Latin hexameter and pentameter. AJPh 51, 358-371.

Description of the behavior of the consonant sequence -s # # n- in the first 5,000 lines in Ovidius' Metamorphoseon libri.

247. WILLIAMS, C.B. (1970), Style and vocabulary. Numerical studies. London: Charles Griffin.

Sequence of d/s in 15 Latin poets; data from Longfellow's Evangeline are added for comparison (chapt. 7).

248. WOHL, B. (1974), The phenomenon of Sperrung in Tibullus' Elegies. TAPA 104, 385-428.

Tables of adjective-noun Sperrung, the amount of non-separations, and their respective percentages,
- inversion of the position of the adjective and the attributive genitive,
- Sperrung in the hexameter compared with the pentameter.

249. YOUNG, A.M. (1931), Schematized word-order in Virgil. The Classical Journal 27, 515-522.

Tables of types of word placement according to a schema (ABCAB, ABCBA, ABACB, and so on), the same letters indicating the words which modify each other and numbers indicating the average number of lines required to provide one example of each word sequence.

List of Contributors

Altmann, Gabriel, Ruhr-Universität Bochum, Sprachwissenschaftliches Institut, Postfach 102148, D-4630 Bochum 1, West Germany

Clayman, Dee Lesser, Brooklyn College, City University of New York, Dept. of Classics, Bedford Avenue and Avenue H, Brooklyn, New York 11210, U.S.A.

Gérard, Jean, 15, Bd. Saint-Germain, 75005 Paris, France

Greenberg, Nathan A., Oberlin College, Department of Classics, Oberlin, Ohio 44074, U.S.A.

Grotjahn, Rüdiger, Ruhr-Universität Bochum, Seminar für Sprachlehrforschung, Postfach 102148, D-4630 Bochum 1, West Germany

Job, Ulrike, Ruhr-Universität Bochum, Romanisches Seminar, Postfach 102148, D-4630 Bochum 1, West Germany

Schmiel, Robert, The University of Calgary, 2920 24 Ave. N.W., Calgary, Canada T2N 1N4

Tordeur, Pol, Avenue Royers 146, B-9600 Renaix, Belgium