ON FINDING FRESH EVIDENCE IN OLD TEXTS: REFLECTIONS ON RESULTS IN COMPUTER-ASSISTED BIBLICAL RESEARCH

DAVID L. MEALAND
UNIVERSITY OF EDINBURGH

In this paper I propose to argue for a greater use of scientific method in the study of literature, while offering some details of computer-assisted analysis of literary texts. The bulk of the article will present results of searches of electronic versions of the texts, and will discuss the methods by which the results were obtained in more detail than has been appropriate in other journal articles. The texts are Greek texts from the New Testament period, but the methods involved in searching them have a wider relevance. In any case the presentation of results, and the discussion of method, will lead on to my more general argument. That argument will maintain that computer-assisted research not only makes available new methods of working and new results, but also provides a much needed spur to re-examine the argumentation used in the humanities. In short I propose to make a plea for a far greater use in the humanities than has often been the case, of a more rigorous scientific style, both in conducting research and in presenting results.

I am well aware that this line of argument is liable to provoke at least two aggrieved responses. One will insist that humanities in general, and biblical study in particular, are essentially literary, impressionistic and subjective, and are concerned with the personal, with values, with freedom and with creativity, and that soulless scientific calculation has no place here. A quite different aggrieved response will insist that the standard of argumentation in the humanities in general, and in biblical studies in particular, can stand comparison with the most demanding and rigorous methods employed in the natural sciences. Neither of these responses has escaped my notice, and I concede a little at once. Yes, there is an essential and important difference between the world of literature and values and that of physics and chemistry, and yes, there are already, and have long been in existence, many examples of fine studies in which careful exploration of fresh evidence is backed by rigorous argumentation. Would that this were so more widely! Both points uphold an important truth, but most of those who follow scholarly writing on literature will have their own list of instances of blatant bias, waywardly subjective appraisals, inferences based on vague generali-

zations, and even instances of fallacious argumentation.[1] These are hard words, but attaining accuracy in thought and expression is not an easy task, and each of us depends on others, as well as on our own vigilance, to discover lacunae and inconsistencies in our reasoning.

One of the main problems which we face here is the relation between the extent of the evidence offered and the character of the conclusion drawn. I am well aware that this is not a simple matter, and that outside deductive logic and mathematics it is often the case, and often inescapably the case, that the evidence available for any given conclusion may raise its confirmation, or tend to disconfirm its contradictory, but rarely, if ever, necessitates the conclusion. In fact we have two well known standards which fall short of deductive certainty. In legal matters we are familiar with the notions of proof beyond reasonable doubt and proof on the balance of probabilities. The difference between the former, more demanding, and the latter, weaker standard has been wisely and skilfully used in recent times by victims who have successfully won a civil case when lacking sufficient evidence to achieve a criminal prosecution. (In some of these civil cases the standard of proof was raised above that of proof on the mere balance of probabilities.) The more technical discussions of this problem tend to be conducted in studies of the philosophy of induction and confirmation, but the study of confirmation theory does have considerable implications for methods of argumentation in the humanities. Though there are some significant differences, there are also many similarities between the criteria which confirm or disconfirm theories in a variety of disciplines.

The analogy with the evaluation of rival hypotheses in science is probably in fact much more valuable than the rather loose appeal to legal standards of proof which is more customary in the humanities. There has been intense discussion of this issue in the philosophy of science. One of the main problems has been differing evaluations of the merits of rival theories, and the exploration of the origin, and possible resolution, of these differences. It is often maintained that the problem of evaluating rival hypotheses does not lend itself to the construction of a simple decision-theoretic model. There are, however, values which preferred hypotheses are widely agreed to exhibit.[2] The values include economy, explanatory power, and above all the capacity to make predictions which are validated by the discovery of

[1] Two interesting and very different studies of loose reasoning in biblical studies are J. Barr, *The semantics of biblical language* (London: Oxford U.P., 1961) and D.A. Carson, *Exegetical fallacies* (Grand Rapids: Baker, 1984).

[2] T.S. Kuhn, *The structure of scientific revolutions*, 2nd edn (Chicago: Univ. of Chicago, 1970), 199; see also ibid., 'Reflections on my critics' in *Criticism and the growth of knowledge*, ed. I. Lakatos and A. Musgrave (Cambridge: C.U.P., 1970), 231–78, esp. 262, W.V. Quine and J.S. Ullian, *The web of belief*, 2nd edn (New York: Random House, 1978), 71, and for historical method C. Hay, 'Historical theory and historical confirmation', *History and Theory*, 19 (1980), 39–57, esp. 50–2.

fresh evidence. This last factor is of great importance in the natural sciences. But this criterion is not as rare in the humanities as some might think. In historical study the discovery of fresh documents, and the production of new data from archaeology, though not so frequent as the production of fresh data in the natural sciences, do yet have a considerable impact. Of course new data from archaeology cannot be produced to order, but then there are problems with doing this with sub-atomic particles, and even more so with finding crucial missing data in palaeontology.

This is not the place for a full scale discussion of evidential probability, but it is the place to argue that computer-assisted research offers a golden opportunity to raise the standard of proof in research in some areas of the humanities, including religious studies, and biblical studies.

THE ROLE OF COMPUTERS IN FINDING FRESH EVIDENCE

Even in subject areas where fresh evidence is rarely available in other ways, computer-assisted research is often able to extract new information from an apparently closed set of data. In the study of literary texts, for instance, there is often very great excitement at the discovery of a new poem, or essay, or letter. To some extent this excitement is one of anticipation of a new work which will excite wonder and invite appreciation. I am not immune to or hostile to this aspect of discovery, nor do I wish to disparage the importance of reading existing texts with fresh eyes and ears. But there is another aspect. A new text or set of texts provides what is often routine in the physical sciences, but much rarer in literary studies: the opportunity to test existing hypotheses or theories against fresh evidence. But we do not necessarily require the discovery of a set of Qumran Scrolls, or Nag Hammadi Gnostic texts, or fresh papyri from Egypt, or the first decipherment of an ancient script, to have fresh evidence. We can, with the aid of computers, search large existing texts more thoroughly than ever before. If the texts have never been indexed, and we have an electronic version, we have something far, far more useful than an index or concordance. We can discover all that such invaluable tools can tell us, and far more. We can look minutely at the structure and syntax of sentences in such texts in a way that could previously only be done either in selected samples, or with extreme labour when larger quantities of text are involved.

In order to exploit these resources, what is needed is a blend of imagination and strict logic. This is, in fact, precisely what the writers of works on scientific method tell us is required in the pursuit of research in the natural sciences. There is an important role for imagination in formulating hypotheses, and in finding criteria to test them, but careful and rigorous evaluation must accompany this. If my examples are mainly drawn from the study of literary texts, that is

because that is what I work with, but I do not doubt that similar moves can be made in other areas of the humanities. We can not only test existing theories about our texts by finding fresh evidence, but we can go through the whole routine of confirming or disconfirming the existing theories, amending them by proposing one or more auxiliary hypotheses, and then, and most crucially, *proceeding to test the auxiliary hypotheses by searching for yet further new evidence.*[3] I apologize if what I am arguing seems in the context of scientific methodology very basic and well appreciated. I do not apologize for pointing out that humanities scholars in general seem not to have appreciated just what the possibilities are here. Told of the release of a newly-discovered set of texts, the well-conditioned reflexes send scholars to the telephone to insist that funds for a copy must be found, and this even in British university libraries devastated by a decade of Thatcher-ite Philistinism. Told that a compact disc (CD-ROM) of electronic texts (E-texts) in their discipline has been released, a few may go and use it, and even then the use is most often just to find more speedily some information that was already being sought. The new technology offers us more than that. But enough of such remonstration, let us turn to a few examples of what can be done in practice, of what can be found, and of how it can be found.

ON FINDING FRESH DATA IN ANCIENT TEXTS

My own training was in classics and in biblical literature, and it is in the study of first century texts and related literature that I am involved, so my first examples will be from this area. One of the most vexed questions in New Testament studies is the tangle of Semitic and Hellenic linguistic, cultural, theological, ethical and ideological features which we find in the familiar texts. It is characteristic of the Hellenistic world in general that we find a considerable mixture of elements from classical Greece and its continuing literary tradition, and from Syrian, Persian, Egyptian, Jewish and Roman features. In artistic representation the paintings on the walls at Dura Europos are a fine example from a slightly later period: scenes from the Hebrew Bible are shown with Moses in Greek dress and characters from the book of Esther wearing Persian trousers. The language and style of the Greek literary texts can be similarly mixed. Yet even so some scholars, who would insist most emphatically that all first century Judaisms are Hellenistic Judaisms, can signally fail to identify clearly what is Semitic and what is Hellenic in our texts. Of course there is a good deal of overlap. It is not unusual to find that certain turns of phrase, certain features of style are shared by more than one culture, found in more than one language. But some features which many, or even

[3] On the importance of further testing of auxiliary hypotheses see I. Lakatos, *The methodology of scientific research programmes* (Cambridge: C.U.P., 1978), 32.

most, New Testament scholars regard as Semitisms can be shown in some cases to be no such thing, or to be the kind of thing one can admittedly find in 'translation Greek', but also readily find in quite ordinary Greek writers.[4]

My suspicions about some of these items had been there for a long time, but gradually became stronger and stronger, as I worked on an article comparing features of the Greek of Acts with the Greek of Hellenistic writers such as Polybius. I had chosen Polybius as a well known historian of the Hellenistic period, comfortably prior in date to the writing of Acts. In fact for this task I chose a group of four major Hellenistic writers, three historians and a geographer to be exact, as comparators for certain features of the style of Acts. Polybius (second century B.C.E.) was my first choice, but the others were Diodorus Siculus, Dionysius of Halicarnassus and Strabo (all first century B.C.E.). In a previous article I had selected a wider group of writers spread over the period from Polybius to Lucian, but this time I wished to focus more on some major Hellenistic writers prior to Acts. Were I beginning a new study I would now, I think, lay more emphasis on Dionysius of Halicarnassus, and less on Strabo, but I would still look very carefully at Polybius and Diodorus. Dionysius is one of the most interesting writers, because he is the turning point between the previous writers of Hellenistic Greek or the higher koine on the one hand, and the Atticizing revival in literary Greek which he began, and which gathered pace after him. In fact I suspect that there is still a lot to discover from a careful comparison of Acts and the writings of Dionysius of Halicarnassus.

In the article in question I simply reported the results of the research, and those who would like more detail of the passages in Hellenistic Greek literature I refer to the text of the article[5] published in 1991. Here I am doing something different, which is to look more closely both at the methods used to find the data, and at the implications of the results for the adoption of a more overtly scientific method in humanities research and in biblical research. Consequently in the earlier piece I gave more examples in Greek in context. Here I discuss fewer examples, with a closer focus on the methods being used. The next two paragraphs discuss some of the technicalities of the method I used, and could be skipped by those interested chiefly in the results of the searches.

[4] On the wider issues here see esp. G.H.R. Horsley, *New documents illustrating early Christianity*, Vol. 5: *Linguistic essays, 1989* (North Ryde, N.S.W.: MacQuarie U.P., 1989) and the works there cited. Horsley is very critical of N. Turner. See N. Turner, *A grammar of New Testament Greek*, Vol. 4: *Style* (Edinburgh: Clark, 1976 – earlier volumes of this work were edited by J.H. Moulton and W.F. Howard). See also R.A. Martin, *Syntax criticism of the Synoptic Gospels*, Studies in the Bible and early Christianity 10 (Lewiston: Mellen, 1987), and E.C. Maloney, *Semitic interference in Marcan syntax*, Society of Biblical Literature Dissertation Series 51 (Chico: Scholars Press, 1981).

[5] D.L. Mealand, 'Hellenistic historians and the style of Acts', *Zeitschrift für neutestamentlichen Wissenschaft*, 82 (1991), 42–66.

In order to conduct the search, I used the TLG compact disc in conjunction with the Ibycus specialized micro-computer. The compact disc in question was produced by the Thesaurus Linguae Graecae of the University of Irvine in California. The Ibycus is a micro-computer which looks like most others, but is specially designed for the searching of literary texts. There are other possibilities which I did not previously mention, but which it is more than appropriate to discuss here. The texts are equally available on tape or on floppy disc (you might need to have the tape downloaded to floppy disc and need several discs for a voluminous author, or use up a lot of hard disc space, so the compact disc is in fact easier, if the relevant hardware can be requisitioned). The texts are stored in a special transliteration scheme. Originally I used a mainframe and a normal text editor. In those early days texts such as Plutarch or Philo which I had acquired on magnetic tape appeared on the mainframe in Roman capital letters interspersed with symbols such as / or \ or = or) or (to represent accents and breathings. But Ibycus or other micro systems will display the correct fully accented Greek characters. Using Ibycus one simply selects first the author or group of authors. I did the latter.

Next one selects the works (I selected all their works), and then the pattern. The term pattern is used because we are not searching for a supposed grammatical root, or for a dictionary entry form of a word, but for a form or set of forms which will capture all the inflected variations of a given word. A search for something like ἀναστάς is one of the easiest to describe; I will give details of a more complex search later. The search pattern does not require any diacritics, in fact it is better not to include these, as they may well change in inflected forms. The results all include the diacritics, so they are not left out at the stage where they are relevant and needed. It is better in this instance not to specify whole words, as we wish to find all the inflected forms of the word. One could choose only αναστ as the pattern, but this will find not only ἀναστάς together with all the inflected forms, accusative, genitive, dative and all the plurals, and the (grammatically) feminine and neuter forms, but too many unwanted words as well. In fact the actual pattern I specified was αναστασ OR αναστ ant, and again it is very convenient that the search routine will find both σ and ς when σ is input in the pattern. (Logical OR is obtained by control-O and is shown on the computer screen as a more complex form of a capital V.) Even with this pattern one will also find the noun ἀνάστασις and its inflected forms, until such time as the system incorporates an operator to signify NOT. But this is a small price to pay, and the now reduced number of unwanted forms are readily discarded at a later stage. A further decision which needs to be made is whether or not to include in the pattern the symbol which stands for the space or punctuation which separates words. Ibycus uses a symbol similar to # for this purpose. By omitting the symbol for word division at the start of the word, one can include the discovery of all the compound forms. This was particularly important in Polybius, where it was very useful to detect διαναστάντων and ἐξαναστάς. (Had I needed to exclude these, I would have specified #αναστασ OR #αναστ ant as the pattern.) I have described the search pattern in some detail as I have not had occasion to do this previously, and as the more complex searches are built up from the simpler units. Nowadays, when searching the TLG, many readers might well prefer to use an alternative to Ibycus, such as a standard IBM compatible micro-computer with suitable software. Amongst the alternative software Lbase and Searcher (now Pharos)[6] are for use on a PC, and Pandora can be used on a Mac. Each of these systems works slightly differently, but the same general principles apply to any computerized searching of literary texts.

[6] The TLG disc is available from the Thesaurus Linguae Graecae, University of Irvine, California. Lbase is produced by John Baima of Silver Mountain Software. Searcher is on the point of being replaced by Pharos, and these products come from Randall Smith, previously of the Dept. of Classics, University of California, Santa Barbara; further information can be obtained from Santa Barbara, or in the U.K. from G.C. Neal, Dept. of Greek and Latin, University of Manchester. Pandora is produced at Harvard.

One of the alleged Semitisms was the construction using some form of the past (aorist) participle ἀναστάς 'having arisen . . .'. This is usually translated 'He (or they) arose and . . . .'. Both in the Septuagint and in Acts one can find examples of phrases such as 'he arose and went' or 'he arose and said'. This had led the unwary to describe the usage as a Semitism. But Polybius also has several instances of people arising and fighting, or arising and speaking, or arising and going (*Hist.* 3.74.1, 35.4.9 and 20.11.7, to offer but a small selection of the examples). Polybius is not alone; I also found examples in Dionysius of Halicarnassus (*Ant. Rom.* 9.3.5, 8.11.2, 7.68.6). The number of examples is quite sufficient to show that the phrase is the kind of Greek one expects to find in the more literary writers of the Hellenistic period. In fact the combination of a series of instances in Polybius with some further examples in the more discriminating Dionysius is particularly significant. At this point, rather than go for yet more numbers, I decided that a further step was needed. But this decision needs further discussion, and again this is a point more relevant to this piece than to the initial publication, where I dealt with the matter rather differently.

The initial search immediately demonstrated that the phrase in question is in no way a primary Semitism. It is *not* a phrase found *only* in translation Greek or heavily Semitized Greek, despite the common opinion of several standard reference works. But whether it is a secondary Semitism is a more tricky issue. If we define a secondary Semitism as a phrase found in Semitized Greek but rare or unusual in other Greek, we then become embroiled in comparative frequencies. To some extent this question was partly resolved by the initial findings. Having found at least twenty-eight examples in the admittedly bulky works of Polybius, and also confirmation of the acceptability of the phrase in Dionysius of Halicarnassus, this did not seem to be the place to try to reinforce the argument with further statistics. In any case very different figures could be used. Should I compare the total number of instances in the Septuagint against the total word count of the Septuagint, and then the total in Polybius against the total word count in Polybius? This would perhaps be relevant if examining some syntactic pattern which was independent of the narrative in each case. But the frequency of use of this expression is at least partially determined by the story being narrated. It is notable that instances of 'X arose and spoke' are far more frequent in Polybius in records of deliberative meetings or their equivalent. Similarly it is hardly surprising that a phrase such as 'X arose and fought' is, as one might expect, found more frequently in accounts of battles. But both phrases are found quite sufficiently to make my case. An even more critical factor is whether or not the phrase can be found in classical Greek. If it can, that would further strengthen the case for the phrase being normal Greek, in that it is not only normal in Greek of the Hellenistic period, but also found in the earlier classical writers. Quite

independently of my work Darryl Palmer of the University of Newcastle, N.S.W. (in work not yet published) has found instances in Andocides, just as I was finding examples in Thucydides and Demosthenes, as well as in other classical authors.

I will return to the question of statistical analysis later. In this case more detailed statistics did not seem to be necessary. In order to disprove the assertion that a phrase was a Semitism, my computerized search of a set of electronic texts found enough examples in Hellenistic and classical Greek to be conclusive. I do not doubt that there are Semitisms in the Greek of the New Testament, but this phrase is not one of them.

Two further examples are of interest, the one more for the method of searching, the other more for the question of the force of arguments based on frequency. The first is the search for the syntactic construction ἐν τῷ with the infinitive. The construction is known in classical Greek, but is argued by some (such as N. Turner) to be a Hebraism and non-classical 'in its temporal sense'. In other words Turner and others claim that the phrase is a Semitism when it means 'While x or y was doing z'. Others again had contested the point with isolated examples.

Here the first problem is to find a fairly complete set of examples. It is easy enough to search for ἐν τῷ . . ., but that will not find ἐν δὲ τῷ . . . or other similar variations. And also one will scoop up far too many instances of 'in the' followed by a noun, rather than the infinitive of a verb. The Ibycus system was improved at about the time that this search needed to be made, and one of the changes was critical to the success of this search. Originally the operator for logical AND had been set to include all examples of pattern *a* AND pattern *b* at a distance of up to 200 characters apart. But a new release of the software allowed this to be adjusted, and within the match options one could specify the maximum distance for the capturing of two or three components in the search pattern. So the pattern had to be set with the AND operator throttled back to something like twenty-five characters (individual letters and spaces). Sometimes trial and error is the only way to resolve the exact decision. (Logical AND is displayed on the screen as a more complex version of an inverted V.) Set the operator with too wide a span, and one captures masses of irrelevant text; set it too narrow, and vital examples are missed. However, the sheer speed of the actual searches means that this need not be a problem. It is possible to perform each search of a group of four voluminous authors in less than five minutes. If a given pattern finds too many irrelevant words or too few relevant ones, the search can be repeated with slight variation until the searcher has found all the relevant words and as little irrelevant detail as possible. This does take some time of course.

It is very easy for those new to this kind of research to be unduly

impressed by the initial speed of searching, and not to recognize the importance here, as elsewhere, of the necessity for patient and careful scholarly analysis. The aim must be to extract from the lengthy texts all, and only, those phrases relevant to the point under investigation. It is an illusion to think that computer searches can all be simplified. Of course if one has a tagged text which will identify all infinitives, this task would be relatively easy. But though the biblical texts have been so tagged, there are huge numbers of other Greek texts which have never been morphologically tagged. (There is more than one system capable of doing this, and the article in this issue by T. Bergren discusses just this topic in relation to the morphological tagging of the Septuagint.)

Controlling the AND operator was the first step. The second was producing a list of all probable infinitive endings. The obvious ones are ειν# σαι# ναι# σθαι#, and these were duly specified and found. (The end of word symbol # is very useful here in limiting the search and excluding many unwanted patterns.) But I knew that I had not found all the relevant instances, as I had not yet matched the total of F.H. Allen, who had studied the use of the infinitive in Polybius in a Chicago thesis of 1907 long before the delights of searching Polybius at breakneck speed by computer. (A salutary lesson both to those who think that computer searching is easy or always complete, and to those who think that our predecessors were not thorough. Allen's work had suffered undeserved neglect in some quarters.) Some more thought suggested adding endings such as αν# and ην# which produced examples of the phrase with ἐν τῷ ζῆν in Polybius and ἐν τῷ . . . ὁρᾶν in Thucydides. Anyone conducting computer searches of Greek is well advised to have a copy of a large lexicon and of a very detailed grammar constantly to hand, and the same is probably true for similar work in other languages.

In this part of the search I was able to find some more examples of the phrase in Greek independent of the Septuagint, and prior to the Septuagint, and further to argue from careful reading of the contexts that several of these passages could well have a temporal sense. In this case the search provided some nice examples of how to find a slightly more elusive pattern, and a result which disconfirms the view that the phrase in question is a primary Semitism.

The examples have so far focussed chiefly on linguistic matters. But fresh evidence can be found in other ways. In looking at the closing verses of Acts I ran searches on three of the key words in the final two verses.[7] The word μίσθωμα is clearly used of rent in civil contexts, though LSJM and, to a lesser extent, Bauer persist in offering for Acts 28.30 a meaning found nowhere else. In that context (as in some other contexts in Hellenistic Greek) ἀκωλύτως

[7] D.L. Mealand, 'The close of Acts and its Hellenistic vocabulary', *New Testament Studies*, 36 (1990), 583–97.

(unhindered) has theological overtones, and the study of these throws light on the redactional significance of Acts 28.31. But ἀκωλύτως is also regularly used in rental contracts, as a search of the papyri with Ibycus showed. This suggested that διετία (two years) might have similar associations, and a further search confirmed this. It is therefore possible to turn up new evidence which not only bears on the philological aspects of the language, but also on its semantics. The combination of the similar associations of the three terms suggests that the author of Acts may have adapted a tradition about Paul renting a place to live in Rome unhindered for two years – standard conditions for one of the standard periods. That the author of the finished literary text saw theological significance in the lack of restraint on his leading character need not surprise us, and some of the fresh evidence bears on that also. Some of the Hellenistic texts speak of the unhindered activity of fate or the gods.

ON FINDING FRESH DATA WHEN USING VOCABULARY TESTS

So far I have discussed three examples, one involving a fairly simple search and a very conclusive result, the other still giving a clear result, but involving a slightly more complex search, the third involving questions of semantic range. None of these has involved stylometry. I have deliberately delayed discussing stylometry, in order to focus on examples in which other stylistic features get more prominence. But a lot of effort in computer-assisted research has gone into stylometry, and I do propose to discuss it, but I do wish to insist that it is by no means the only way to conduct computer-assisted literary research.

Certain authors tend to attract the attention of stylometrists. Shakespeare is one, Plato another, and Paul a third. In recent years we have seen full-length studies[8] by Tom Horton on Shakespeare, by Brandwood and by Ledger on Plato, and by A.J. Kenny and by K.J. Neumann on Paul, as well as continuing contributions from A.Q. Morton and others. The stylometric equivalent of the search for the Holy Grail is the development of a method which will achieve the assignment to the correct author of very short pieces of text. In practice, to achieve this reliably for samples as short as 750 to 1,000 words is still contentious, though some of the recent studies have achieved significant levels of accuracy in assigning test samples.

What is quite extraordinary in New Testament studies is that there is generally very considerable reserve about large scale stylo-

[8] T.B. Horton, 'The effectiveness of the stylometry of function words in discriminating between Shakespeare and Fletcher' (Diss. 1987, Dept. of Computer Science, University of Edinburgh), G.R. Ledger, *Re-counting Plato: a computer analysis of Plato's style* (Oxford: Clarendon Press 1989), L. Brandwood, *The chronology of Plato's dialogues* (Cambridge: C.U.P., 1990), A.J. Kenny, *A stylometric study of the New Testament* (Oxford: Clarendon Press 1986), K.J. Neumann, *The authenticity of the Pauline Epistles in the light of stylostatistical analysis*, SBL Diss. Series 120 (Atlanta: Scholars Press, 1990).

metric studies, yet astonishingly wide acceptance of grossly oversimplified vocabulary tests on short pieces of text, for which remarkably little statistical justification has been offered. I propose first to offer some words of caution about the latter practice.

A regular area for authorship assignment on the basis of a few brief notes about vocabulary is in Synoptic studies. In Luke, for instance, it is quite common to decide that a few verses are either redactional and due to the author, or pre-Lukan and to be assigned to the tradition, on the basis of a few words. I must admit to having done this myself in the more distant past, and to still giving some weight to the practice, while having very grave worries about the fact that its statistical basis is extremely ill defined. I am not therefore attempting to discredit anyone here, but rather to voice a considerable unease about the widespread use of a method which needs more rigorous scrutiny. There is some acute comment on analyses of Markan vocabulary in C.C. Black's recent work.[9] Even in the analysis of Luke there are problems. There, at least, we can draw inferences from differences between Luke and Mark on the one hand, and from similarities between Luke and Acts on the other. But even then the fact that a given word is inserted by Luke in a passage derived from Mark, and also used by the author in Acts, is not a guarantee that all passages containing such a word are due to the redaction by Luke. In fact arguments are usually based on the detection, in a given set of a few lines, of several words found regularly either in Acts, or in Lukan additions to Markan material. But even so each item has to be scrutinized. I cite an example.

In a passage found only in Luke the author reports that Jesus was sent by Pilate to be interrogated by Herod Antipas. Is Luke 23.6ff from a source or redactional? A decision on the stylistic issue has important implications for deciding whether the story is a Lukan creation or rests on older tradition of some kind.[10] One of the

[9] C. Clifton Black, *The disciples according to Mark*, Journal for the Study of the New Testament Supp. 27 (Sheffield: Sheffield Academic Press, 1989), 205–12, whose comments on p. 209 on the lack of statistical basis should be noted carefully. There is, however, a very brief example of a chi squared test on Lukan vocabulary in G. Herdan, *The calculus of linguistic observations* ('s-Gravenhage: Mouton, 1962), 34, but he is using it as an illustration of a calculation, and he limits himself to the conclusion that the words regarded by Morgenthaler as favourite expressions of Luke are 'not, on the whole, also favourite expressions of Matthew and Mark' Herdan hints at a good way to proceed, but did not himself proceed far along it in work on the Synoptics.

[10] The issues are summarized in major commentaries such as J.A. Fitzmyer, *The Gospel according to Luke* (New York: Doubleday, 1985), 1479; see also I.H. Marshall, *The Gospel of Luke* (Exeter: Paternoster, 1978), 855 with less emphasis on the argument from vocabulary. Fitzmyer cites J.C. Hawkins, *Horae synopticae*, 2nd edn (Oxford: Clarendon Press, 1909). Hawkins's rules list words which are used at least four times in Luke, and also are either not found at all in Matthew and Mark, or are found in Luke at least twice as often as in Matthew and Mark together. Hawkins and others also list further words which fall slightly short of this requirement, but are held to be 'more or less characteristic of Luke'. Further brief discussion of this issue can be found in Fitzmyer, *Luke*, 109–12. Careful study of pre-Lukan vocabulary can be found in H.

arguments deployed is that the verb 'to send over/back' ἀναπέμπειν is, apart from one passage in Paul, used only in the Lukan writings in the New Testament. But the word appears only five times in total in the New Testament. Three of the uses are in the passage under examination in Luke, there is one usage in Acts, and one in Paul. The word is appropriate to the subject matter in Luke 23. I do not doubt that the word may have been introduced by the author of Luke, but I do not think that the data offered actually establish this. The figures are just far too low for reliable analysis without further argument to support the case. Two further instances are slightly better supported: in the same passage we find ἱκανός (several, some, or many) and σύν (with). The former of these is found three times in Matthew, three times in Mark, seven times in Paul, nine times in Luke and eighteen times in Acts. Clearly the author of Luke-Acts uses the word and in the relevant sense, but we also have instances of others also using the word, and in some instances with the relevant sense. I must admit that it seems intuitively plausible that the word is more likely to have come from author than from source, but the usage in Mark 10.46 shows us that the word could have appeared in a Lukan source. In the case of σύν the figures for Luke-Acts are much higher, though again there are a few instances in each of the three other Gospels, in James and 2 Peter, and many in Paul. I cite the examples above, to show that there is widespread acceptance of simple stylometric claims based on very limited samples of vocabulary whose statistical basis has, as far as I am aware, never been adequately tested. Schürmann often offers careful argumentation about pre-Lukan vocabulary in the work cited above, but his methods do not include an analysis of statistical significance.

In order to explore this problem further, I tried the following empirical test. I took a piece of text at random where Luke and Mark run parallel, and which Luke presumably derived from Mark and adapted. The passage was Luke 18.35–39, which contains fifty-five words. On the assumption that the verbal differences between Luke and Mark are due to redactional changes by Luke, I listed ten words and one expression apparently attributable to Luke's redaction. I then checked the ten words, to see how many would have been noted as redactional, if we had used the lists of Lukanisms, and had not had Mark's text for this passage. This test is inevitably somewhat biased in favour of the method I am wishing to criticize. The lists are derived from various comparisons, which include comparisons between Mark and Luke, and which

Schürmann, 'Protolukanische Spracheigentumlichkeiten?' in *Traditionsgeschichtliche Untersuchungen zu den synoptischen Evangelien* (Düsselforf: Patmos, 1968), 209–27.

Further studies of Lukan vocabulary can be found in works such as B.S. Easton, *The Gospel according to St. Luke* (Edinburgh: 1926), F. Rehkopf, *Die lukanische Sonderquelle*. Wissenschaftliche Untersuchungen zum Neuen Testament 5 (Tubingen: 1959), J. Jeremias, *Die Sprache des Lukasevangeliums* (Göttingen: Vandenhoek & Ruprecht 1980), *Synoptic studies*, ed. C.M. Tuckett, JSNT Supp. 7 (Sheffield: 1984), J. Dawsey, *The Lukan voice* (Macon: Mercer, 1986), M.L. Soards, *The passion according to Luke*, JSNT Supp. 14 (Sheffield: 1987).

include reference to this passage, but the effect of the bias is limited. If we suppose that we had all the rest of Mark, but that the equivalent passage in Mark 10.46–48 was damaged and unreadable in all our texts of Mark, how many of the ten redactional words would we identify as redactional by the kind of test used by the commentators on Luke 23? The answer seems to be four. There are six further words which I am assuming to be redactional from actually reading Mark 10.46–48, but which would not be picked out by the lists. Three of these six would be selected if the rules were relaxed even further than those used in Hawkins' subsidiary lists. In these three instances each of these words is used more often by Luke than by Mark or Matthew, but not much more, and less than five times in all. This might suggest that the rules derived from Hawkins and his successors are in fact too strict. But we need to bear in mind the following further obvious, but very important, consideration. How often would the rules identify something as redactional that is in fact due to the source?

Ideally one would like to find a passage in Luke which has a Marcan parallel containing two or more of the words we would, on the vocabulary tests in question, otherwise have assigned to the Lukan redaction. This is a severe test, though if we can find such a passage it would certainly severely challenge the method under scrutiny. What I have done is this. I took a sample of five of the words said to be 'characteristic of Luke' (ἀναστάς, ἐρωτάω, κλαίω, νῦν, σύν). I then looked in Mark for passages in which two of these words appeared in reasonably close proximity to one another. The theory being tested rests on the fact that the words in question are found at least twice as often in Luke as in Mark. In fact none of these words is found more than six times in Mark. My tests are hardly exhaustive, but I did find three passages each containing a different pair of these words: Mark 4.10, 7.24–26, 15.27–32. Had we found these passages in Luke we might well have assigned these words to his redaction, and done so mistakenly. The theory admits that a few of these words do appear in Mark, but I was surprised to find so readily passages where pairs of the words appeared within a few lines of one another. Whether or not I should have been surprised could probably be assessed by collecting more data, and making a series of calculations of vocabulary frequency. Luke has in fact no equivalent to Mark 7.24–26, and in the other two passages preserves only one of the four of his 'characteristic' words served up to him on a platter by his source, Mark! This is in itself rather revealing. I do not claim to have solved this problem, but I do think I have done enough to show that some of the vocabulary tests, widely accepted by many major New Testament scholars, need some much more careful scrutiny.

The examples I have taken are derived from Mark. The reason for this is that the vast majority of New Testament experts would agree that

Luke used Mark. If, therefore, we can find supposedly Lukan words grouped in a Markan passage, we can show that the occurrence of such words in Luke is due to a source rather than the redactor in at least some instances. I am not denying that Luke uses these words more frequently than his sources, but rather I am claiming that their appearance in Luke can sometimes be due to his sources. If I can show this for Luke in relation to Mark, when Hawkins' lists have allowed for the evidence of Mark, it is very likely that the lists of words will be even less reliable in distinguishing Luke from his special source known as L. We have Mark, we do not have that special source.

I made some further tests. I took the lists of Lukan words given by Hawkins and Fitzmyer, and searched for further combinations of these in Mark. In the story of the demoniac at Mark 5.9–14 I found in six verses four instances of 'Lukan' vocabulary (ὄνομα (2), πέμψον, γεγονός). Luke kept two of these. In Mark 5.29–39 I found four instances of 'Lukan' vocabulary of which Luke kept three. In the story of the arrest in Gethsemane in Mark 14.38–49 I found five 'Lukan' words (πειρασμόν, παραγίνεται, ἀφεῖλεν, συλλαβεῖν, καθ' ἡμέραν), of which Luke kept three. I do not dispute that Luke is normally more 'Lukan' than Mark. But in these specific details Mark is, it seems, more 'Lukan' than Luke.

The tests outlined above go some way to setting up ways of confirming, or disconfirming, by empirical means, the sort of vocabulary analysis widely accepted at present. (I use the term 'disconfirm' in the technical sense meaning 'show to be less likely'). The idea of conducting this kind of test arose, in fact, when I was attempting to analyse a passage in Josephus' *Antiquities* for a longer study which is still in progress. I am, however, able to offer an interim report here. Few passages in Josephus have commanded as much scholarly attention as this one. The text appears to offer a biographical notice of Jesus of Nazareth written by the Jewish historian Josephus before the end of the first century of our era. It is therefore, for many scholars, one of the most important sections in the whole of the writings of Josephus. This fascinating passage comprises a paragraph of about one hundred words in a text several times longer than the New Testament. We have therefore one of the ingredients for a stylometric study, but not some of the others. We have a very large sample of undisputed text, but only one very small sample of doubtful text. This is a better situation to be in than the one just described, but it does not satisfy the full requirements for the kind of stylometric exercise usually undertaken. These tests are usually conducted on samples of text of about 1,000 words or sometimes 750 words. Either we are to abandon any attempt to form a judgement on the authorship of this highly controversial section of text, or we have to try to devise tests that will give us some rational grounds for reaching a decision,

even if they are based on a different type of calculation from that more commonly used in many stylometric studies.

So what method are we to use when faced with this problem, where the disputed sample is a paragraph of slightly less than one hundred words? Can such a method be found, and found to be reliable? How are we to go about it? Though this paragraph is a very short sample, it is at least found in a body of writings amounting to somewhere near half a million words. We are therefore not short of material which is genuinely part of the output of Josephus (even if he may have employed assistants to improve his Greek style).

Is this highly controversial passage a genuine part of Josephus' work? Is it a forgery in its entirety? Or is it partly genuine and partly tampered with, and if so which parts might be genuine? The many studies of this passage have been of very varied levels of sophistication and reached very varied conclusions.[11] Some of the phrases look suspiciously like Christian adaptations, and some scholars have indeed suspected the entire passage to be a forgery. If all, or even most, of the passage is genuine, it would provide one of the very few pieces of independent evidence about Jesus, written by a near contemporary.

A number of studies have examined the words and phrases used in the passage. Since the concordance to Josephus was published, these studies have increased in detail. More detail still is possible using an electronic text. But a further very important point is the use of controls. We need to ask not only if a given turn of phrase in this passage can be found elsewhere in Josephus, but whether it might be a common phrase found also, say, in Philo or elsewhere. If it is, then, though the phrase may be matched in Josephus, it is nonetheless a phrase that several writers might have constructed. Differentiating such phrases from ones that are peculiarly Josephan in style is important. But we also need to ask, at least in relation to key phrases, whether they match the style and language of early Christian texts. A particular suspect is Eusebius; so early Christian texts prior to, and especially including, the works of Eusebius need to be checked. In order to do this I looked not just at a few phrases to see if they are matched in the New Testament, but at several of the phrases to see if they are matched in any of the early Christian writings (up to the time of Eusebius) that are available on the TLG disc.

Even having done this, there is a problem of evaluating the results. We are still in the early stages of developing a reliable and well established set of criteria for assigning authorship, even in the case of 750 word samples, though a lot of progress is being made on this problem. Yet some of the passages we most wish to test are far shorter

[11] The literature is massive and cannot be listed here. One of the best of the recent studies is P. Bilde, 'Josefus' beretning om Jesus', *Dansk Teologisk Tidsskrift*, 44 (1981), 99–135; one of the most recent with ample bibliography is J.P Meier, 'Jesus in Josephus: a modest proposal', *Catholic Biblical Quarterly*, 52 (1990), 76–103.

than that, as this one is. There are vocabulary tests widely used in Synoptic studies, but their statistical basis is extremely slender, as I have argued above.

In analysing the Testimonium Flavianum, as this passage is often styled, I decided to make an attempt at an initial pre-statistical appraisal by using three scales. These are: a) frequency of such a phrase elsewhere in Josephus, b) how common the turn of phrase is in Greek literature more generally and in early Christian texts, and c) how complex the phrase is. While this attempts to capture some of the variables, it has to be admitted that a considerable element of judgement remains, and the process is by no means a matter of simple calculation. On scale A, for instance, I rated each phrase as follows. If it appears nowhere else in Josephus 0. If it appears in one other place 1. If it appears two to four times elsewhere in Josephus 2. If it appears more frequently 3. On scale B I rated each phrase in the following manner. If it is very common in Greek writers 0, if I found or expected to find a good number of examples 1, if the phrase can be matched only rarely outside Josephus 2, if very rarely outside Josephus 3. On scale C I tried to calibrate the complexity of the match between the phrase in the disputed passage and similar phrases elsewhere in Josephus. A match of five words out of six would score 3, of a pair of words 2, of a single but slightly unusual word 1. One has to exercise judgement here, and give little weight to very common words, and more to less common words. I do not claim that this system is fully mathematical; it clearly is not. But it does attempt to give a provisional weighting to some of the judgements that can be made, and which can be defended with numeric evidence to a certain extent. If we now total the results from these three scales of 0–3, then the final results are plotted on a scale of 0–9. This is one way, however crude, of combining the factors into a single scale.

At this point one further control is employed. The process is repeated on an undisputed sample of Josephus. This gives us some indication of the point on the scale 0–9 at which we would place a genuine piece of Josephus' writing. The results of all the above work then allow us to draw a conclusion about the genuineness of the passage based on a more rigorous procedure than hitherto. I am well aware that the system I have adopted in investigating this passage is by no means a standard statistical one backed by widespread studies. On the other hand I am quite sure that it includes far more safeguards than I have found in those of the very many previous studies of this passage that I have consulted.

It is possible to consider further controls. The method could be repeated on passages from Philo and from early Christian writers, to see what the likelihood is that a passage from another writer might be wrongly attributed to Josephus. Of course we would have to choose passages which were as near as possible to the kind of thing that Josephus wrote. And it is also, of course, the case that such a test

would not eliminate the possibility of a skilful forger combing Josephus for suitable phrases, and using them in order to make his interpolation look more like the genuine thing. The theory behind the general run of stylometric work is that one needs large samples of text to detect traits which are less likely to be consciously contrived. This is often done by examining function words such as 'and', 'but', 'it' and the like. But that method is not open to us on this sample due to its brevity.

Work of this kind can be very time consuming, and also requires far more space to present in detail than is possible here. In this article I am concentrating on the methods used, and the reasons for them, and arguing a case for a general tightening of methods of argumentation, and the adoption of a more scientific method of procedure. Building in the controls is an important part of this. It is not sufficient simply to look for matches between phrases in this passage on the one hand, and elsewhere in Josephus on the other, or even to see if there is something comparable in the New Testament. A wider range of controls is needed: Philo as an example of a second Jewish writer of Greek on the one hand, and a wide range of Christian texts on the other. With these controls we are more likely to detect what is peculiar to Josephus, what is general in Greek, especially in the Greek of Jewish writers, and what could well have been written by an early Christian.

In the passage under examination my provisional conclusions are that of some 23 turns of phrase 9 are ranked at 5, 9 at 4, 4 at 3 and one at 2. For the control sample from Josephus, I selected the beginning of the account of the conversion to Judaism of Queen Helena. In this passage I selected six phrases. (These tests can be very time consuming.) Of these six phrases I ranked three at 5 and three at 4.

Consultation with a statistician confirmed that though the passage under examination is very short, good use can be made of the very large quantities of undisputed text available. We have several blocks of text each over 300,000 words in length: the Septuagint, Philo, Josephus himself, the New Testament with early Christian writings up to Justin, Clement of Alexandria and Irenaeus, Origen (who weighs in at something like a million words) and Eusebius. All of these blocks of text are very substantial and some are massive. This fact makes it likely that it will at the end of the day be possible to apply a rather more rigorous statistical procedure to the results. The fact that many of the phrases in the disputed passage are very similar to other phrases in Josephus is likely to be significant. It is also likely to be significant that many of the phrases are not much matched in early Christian writings. Where phrases are sometimes matched in Clement of Alexandria, Origen or Eusebius, it is rarely the case that they are matched in these writers and not matched in Josephus, or matched in these writers proportionately more than they are in Josephus. But this analysis is still in progress, and, when published, will need to be filled out with much more numerical and stylistic detail.

Though further work remains to be done, a sufficiently clear picture is already beginning to emerge. When I began to study the passage about Jesus, I had initially thought that little if any of it would turn out to be in the style of Josephus. But my results so far actually suggest that on the more extensive tests, provisionally reported here, the bulk of the passage about Jesus in Josephus is genuine.[12]

ON FINDING FRESH DATA IN STYLOMETRIC ANALYSIS

Where longer passages are the subject of investigation, ever more refined methods are being developed. Unfortunately these studies tend to be greeted with a deafening silence by scholars of a more conventional literary bent. This is a great pity, as the methods are often carefully tested in a way which contrasts sharply with the simplistic vocabulary tests I have argued to be too readily accepted. Several recent studies have made use of various multivariate statistical analyses. In the case of Ledger's work on Plato these analyses were based on criteria derived from letter frequencies in 1,000 word samples. The criteria are restricted in this way, but a very wide range of tests was run. Despite questions about some of these tests, the overall results are impressive. They go a long way both to indicating that the criteria do discriminate effectively in assigning most samples from the wider set of samples from works by seven authors to the correct author, and to producing results in Plato's writings that, while not always conforming to prior expectations, are both plausible and at crucial points corroborated. For instance his tests find some of the dialogues to be both genuine, and also in the correct sequence if they are genuine.[13]

Horton's work on Shakespeare and Fletcher also used multivariate analyses, but on a wider range of criteria. He was able to demonstrate a high degree of accuracy in assigning samples derived from texts of undisputed authorship, before working on the disputed samples. Horton's work was on a problem well suited to stylometric analysis in that it satisfied three criteria. (1) The set of likely authors was well defined. (2) There were ample undisputed samples from each of the likely authors. And (3) there were ample sections of disputed text.

More specifically related to New Testament study is Neumann's work on the Pauline Epistles. Here the problem is more intractable in

[12] The phrases I ranked lower on the criteria listed were 'if indeed one ought to call him a man' (only very partial parallels in Josephus and there are parallels in Eusebius), 'and was a teacher of' (limited parallels in Josephus, parallel in Eusebius), ' . . these and other marvellous things' (only partially paralleled in Josephus, similar phrase in Eusebius), 'the tribe of the Christians' (no very close parallel in Josephus, parallels in Eusebius), 'has still to this day not disappeared' (no close parallel in Josephus, Eusebius does have a lot of close parallels to 'still to this day'). I do not assume that all of these phrases are Christian additions, but it does look as though there has been some elaboration. In other cases crucial phrases, such as 'was called' or 'they said that', may well have been omitted from phrases whose style is Josephan, but whose present content suspiciously close to Christian claims. I intend to publish the full results in much more detail in due course.

[13] See note 8 above.

that the set of possible authors is not well defined, and the number of samples is limited. Again multivariate analyses were used, and here also a wide variety of criteria were tested on a set of passages from four early Christian authors (Paul, Hebrews, 1 Clement and Ignatius).[14] The four or five best discriminators were selected from the criteria, and these were then used to explore a further set of passages from the four authors, and from the dubia on the fringe of the Pauline corpus. The criteria eventually selected from a very wide and varied range included word length, the use of relative and indefinite pronouns, words with initial tau, and the position of the first noun in (modified full stop) sentences. By using a series of samples each of which contained 750 words carefully purged of quotations, Neumann was able to pursue clearly defined statistical procedures a long way. (Particularly noteworthy is his rehabilitation of word length as one very important criterion at least in multivariate analysis.) In his conclusions he occasionally puts forward auxiliary hypotheses without further testing. For instance he finds Ephesians Pauline, despite samples failing the tests based on initial tau. The samples are three or more times the standard deviation above the Pauline norm. He also concludes that Colossians and 2 Thessalonians are Pauline despite their being around three times the s.d. above the Pauline mean for relatives and indefinites, and 2 Thessalonians being 3.9 times the s.d. above the Pauline mean in delaying the first noun. In other respects Neumann's work clearly shows great care and resourcefulness. When he obtains unexpected results, he carefully rechecks samples, or tries variations in method. Of course the reader can always ask for more, but the researcher has to decide to conclude somewhere and to publish, and Neumann has clearly done a great deal. That there is more still to be done he freely acknowledges, and it is the sign of a fruitful approach that it leaves the reader with several ideas for yet further tests. In scientific work it is always necessary to acknowledge predecessors, it is also important to signal where one believes there is further work yet to be done.

In an earlier article I had used univariate chi squared tests to reassess some of Morton's claims about positional stylometry. I concluded that tests on Morton's key words such as 'and', 'but', 'for' and 'if' as first or second words in sentences in Paul did not require us to limit the Pauline corpus to as few as four letters, though they do suggest that not more than seven letters are genuine. Neumann's work confirms and extends the doubts I expressed about the reliability of some of the criteria, and in fact indicates that his criteria are preferable to those proposed by Morton, in that they vary less within the different groups of letters, and more between them. Of course Neumann is chiefly using multivariate methods, and that is a very important difference. The process of identifying criteria with a

[14] See note 8 above

low within-group and a high between-group variation is very important. Neumann's work also points to the conclusion that the probable total for Pauline epistles is greater than four but not more than ten. He includes Ephesians, Colossians and 2 Thessalonians, but has to admit difficulty in doing so. His results clearly suggest another author for the Pastorals.

Neumann's work is very significant for its use of samples from other epistolary literature, for starting with a wide range of possible criteria, and for using well recognized statistical procedures such as Stepwise Discriminant Analysis and Canonical Discriminant Analysis (known to SAS users by names such as Stepdisc, Candisc, and Discrim). The concluding tables are clear to anyone who can do simple arithmetic involving means and standard deviations. (It is a pity that the published version does not include an appendix giving the full totals for each sample, though many details are included.)[15] Publishers do tend to be resistant to including long appendices giving essential data. Scholars will either have to try harder to insist on their inclusion, or ensure that the data are carefully preserved and made available to future researchers. Ledger, for instance, very correctly took the latter course when the former was not available.

The works discussed above represent notable advances in the field of literary statistics. It does seem that considerable progress has now been made in discriminating between sets of 750 word samples by different authors. But this of course still leaves one with a considerable problem when faced with many disputed passages that are far shorter than this sample size. Even here there may be possibilities deriving from work by N. Timmins on syntactic pattern recognition. An initial report on this by Timmins is currently under consideration by one of the journals for New Testament studies.

The use of electronic texts makes possible the acquisition of a whole range of additional data even from texts that have long been studied. The possibility of searching these texts by computer[16] allows and

[15] As well as the reservations expressed in the text, I have some further concern about the fact that the samples at first also included text from Philo, Josephus and Epictetus, but when these writers were included the discrimination between the Christian writers was less clear, and several samples were misclassified. Neumann is commendably candid about this, and discusses the implications with care.

[16] I have not included here discussion of the use of an editor to manipulate the electronic text or the output from it. In addition to the well known multi-lingual word processors now widely available, there is a lot to be said for continuing to use powerful line editors which allow complex repetitive tasks to be performed. These can make short work of marking the ends of lines, inserting a space at the start of each line, counting every instance of a given pattern, checking the length of a text and so forth. There are also programmes such as TACT and the Literary Apprentice which are designed for the analysis of literary texts. TACT is available from the University of Toronto; the Literary Apprentice is being developed at Harvard. It is also wise, where possible, to double check at least some samples with different texts on different systems.

indeed lends itself to the discovery of fresh evidence. It is important that the new possibilities are harnessed to a very careful assessment of the argumentative structure of literary analysis. There is a distinct possibility of closing some of the gap between work on literary texts and work in the sciences, but if we are to attempt to do this it must be done in a way appropriate to the study of literature. In order to achieve this aim, one of the most important requirements is to formulate carefully hypotheses which can be confirmed or disconfirmed by the subsequent search for the relevant evidence. I hope not to have laboured this point, but rather to have brought it to the surface explicitly from time to time, and to have offered other examples where it is implicit in the procedures being followed.

ABOUT THE AUTHOR

DAVID L. MEALAND is Senior Lecturer in the Department of New Testament in the University of Edinburgh. He is currently interested in the Greek style of writings such as Luke-Acts in comparison with other Greek writings of the Hellenistic period such as the works of Polybius and Dionysius of Halicarnassus. He has published articles on this area in *New Testament Studies, Zeitschrift für die Neutestamentlichen Wissenschaft* and *Journal for the Study of the New Testament.*

CONTACT ADDRESSES

Post: Dr David L. Mealand, Faculty of Divinity, University of Edinburgh, New College, Mound Place, Edinburgh, EH1 2LX
Phone: (0)31–225 8400
E-mail: D.MEALAND @ UK.AC.EDINBURGH
Fax: (0)31–220 0952