

## NOTES ON COMPUTING

### *Report on the GRAMCORD Project*

Project GRAMCORD (an abbreviation of GRAMmatical conCORDance) began in 1976 as a dream of Dr. James A. Boyer, then chairman of the department of NT and Greek at Grace Theological Seminary. Paul Miller, then a research student at the University of Indiana in Bloomington, worked with Dr. Boyer to produce the first version of a tagged text of the Greek NT (the Project began with a transliterated form of UBS<sup>2</sup>) and of a retrieval package program. In 1979, the Project moved to Trinity Evangelical Divinity School, under whose auspices it has continued to mature and develop. In the Spring of 1986, the Project was placed under the direct control of The GRAMCORD Institute, a not-for-profit educational and research corporation indirectly controlled and still largely financed by Trinity. The Institute is developing an array of software related to biblical study, including the adaptation of the GRAMCORD system to biblical Hebrew.

This report, however, focuses more narrowly on the NT GRAMCORD system. The technical details—programming language, size of text files and program files, and the like—have been discussed elsewhere, most recently in the excellent and informative book by John J. Hughes (*Bits, Bytes, and Biblical Studies: A Resource Guide for the Use of Computers in Biblical and Classical Studies* [Grand Rapids: Zondervan, 1987]). The aim here is to provide both a simple description of GRAMCORD for those who know nothing about it, and an update for those who know a little of the Project and would like to learn of its current status.

The database is a tagged version of NA<sup>26</sup>, in transliteration (and therefore without breathings or accents) and capital letters. "Tagged" in this context means that each word's morphological characteristics are recorded, along with (in many instances) various functional indicators. For instance, *anthrōpos* will be recorded as ANQRWPOS, and tagged as a nominative, masculine, singular noun; but words that are not inflected (e.g. *kai*, *hina*) are nevertheless tagged according to certain functional specification. Thus *kai* (appearing as KAI) may function as an adverb, a correlative coordinating conjunction, or a copulative coordinating conjunction, and will be tagged appropriately in each instance.

Although there are now other tagged New Testaments (especially the Friberg database, based on UBS<sup>3</sup>), the power of the GRAMCORD system lies in the package program that enables the user to retrieve a very wide range of constructions. Of course, GRAMCORD can be used to retrieve every instance of a particular word, but normally that is an inefficient use of the system when standard concordances provide all the relevant information. Even here, however, GRAMCORD may prove helpful in pulling out all instances of some subset of the occurrences of a particular word—e.g. not every instance of *en*, but every instance of *en* plus the articular infinitive. In this instance, of course, it must be admitted that such a list could be compiled by reading through the relevant article in Moulton and Geden, with its own helpful tagging system; but it takes little imagination to think of subsets of words that are not tagged in any of the normal word concordances.

Beyond words, GRAMCORD can very easily retrieve every example of any particular morphological form. For instance, it is a simple matter to display every instance of a perfect passive participle, or of an aorist infinitive. Where GRAMCORD displays its greatest power, however, is in its ability to retrieve any grammatical construction of any complexity or length that is definable in morphological and/or positional terms. Utilizing the package program, the user can write a command file that will retrieve not simply the aorist infinitive, but, say, each instance of the articular aorist infinitive compounded with a double accusative and located within a subordinate clause.

The potential exegetical value of such lists is not hard to imagine. When a biblical interpreter seeks to understand the meaning of a word in a particular context, first recourse is usually made to the standard lexica (which after all are nothing more than the inferences drawn by scholars who have themselves resorted to word concordances), and then to the concordances. The latter will provide exhaustive lists that enable the interpreter to draw important distinctions and refinements. Until now, when the same interpreter sought to understand the semantic force of a particular syntactical unit, recourse could be made only to the grammars. Except for certain constructions that have been exhaustively retrieved by hand, or for others that are so rare their infrequent occurrences in the NT are already well known, the examples used by the grammarians could never be more than a selection of the available evidence. Careful inductive study of syntactical units to the same order of precision that has been theoretically possible for words has therefore been *principally* impossible. GRAMCORD is an attempt to remove this impasse. Exhaustive lists will therefore serve as the essential raw material used by the next generation of grammarians.

Other gains can be promised. Pedagogically, GRAMCORD has the potential for enriching syntactical study amongst advanced students of NT Greek. Just as an older generation of students participated in the joy of discovery as they began to learn how to use word concordances intelligently, so a new generation of students can learn advanced syntax the same way. For several years, advanced students at Trinity Evangelical Divinity School have been required to learn how to handle the GRAMCORD package, to write a suitable retrieval file, and then perform an inductive study of the results. Many of the student papers were mediocre at best, some far too speculative, some quite lacking in scholarly controls or deep analysis—exactly as is the case when students attempt their first inductive word studies. But the best paper or two each year has been publishable, pushing back the frontiers of knowledge and quietly modifying or correcting or refining the existing grammars on the subject. In any case, the deepest value has been to the student who by this exercise has learned a little of how grammars are formed, what induction is all about, what rôle antecedent theories of language play in the inferences that are drawn, and so forth.

Like any tool, this one is not without its dangers. Two or three obvious ones deserve mention at this point. In the first place, the GRAMCORD database provides no textual variants. Where there is deep division of opinion as to the appropriate morphological tag (e.g. *Is hēte* in a particular context indicative or imperative?), the database is flagged and the user is warned in the output; but textual variants themselves are not included, even though in not a few instances they affect not only the interpretation of a particular passage but the statistical weight given to this or that construction. In the book that will issue from the computer package (about which more is said below), this limitation is removed.

A second problem is that the database covers only the NT, and the NT is only a sliver of the broader corpus of Hellenistic literature. One must therefore be rather cautious about extrapolating one's conclusions beyond the NT. But that does not mean such (admittedly preliminary) work is without value—just as a word concordance of the NT is valuable in its own right, even though there is still more value in an inductive study through the broader literature. We have to start somewhere. Plans are afoot to provide similar retrieval power for the LXX, Josephus, a Greek writer uninfluenced by Semitic language or thought (such as Polybius), and some of the papyri. Inevitably, the pace at which such work will proceed will depend on resources both financial and personal.

The third problem area is by far the most difficult. For all but the simplest constructions, writing command files using the GRAMCORD package is a

more difficult and subtle task than many users recognize. Here it may be helpful to say a few words about the nature of GRAMCORD command files. Each file is made up of three sections. The first is the TITLE section. This not only enables the user to label for future reference the construction being sought, but it permits the user to develop a number of subsections in the command file—subsections that look after variables in the targetted construction. For instance, if one were seeking to retrieve every instance of the genitive absolute where the substantival component is a noun, one would have to specify in one subsection of the command file that the noun *precedes* the participle, and in another subsection that the noun *succeeds* the participle. Only in this way will both possibilities be catered for. This control of the subsections of the command file is exercised by the TITLE section. The second is the TARGET section. Here the critical words or parts of speech or morphologically specified elements that the user wants to retrieve are specified. The third and final section is by far the most powerful: it allows the user to specify a wide range of OPTIONS. This includes the number of words that may intervene between any two elements of the target, the specification of the agreements in gender, number, and case that may link two or more elements, the stipulation that punctuation may or may not break up the targetted elements, and much more.

Although these OPTIONS are what make GRAMCORD so powerful, the user must constantly be aware that this is the section of the command file most likely to introduce problems. There are two common sources of error. The first is too little attention to the possible range of examples in the targetted construction. For instance, suppose we search for every instance of *en plus* the articular infinitive compounded with a double accusative. The unwitting user may specify that *en* must be the *first* element in the targetted construction. That will certainly retrieve most examples in the NT; but there are occasional instances where one of the accusatives *precedes* the preposition. Are there any such anomalies where *en* is the preposition? Or suppose that the target is the genitive absolute. How many words may intervene between substantive and participle? Or, to put it more technically, What is the contextfield of the construction? If the user stipulates something too small, what GRAMCORD retrieves will not be exhaustive. But a very large contextfield will certainly pull out alleged instances which on examination are not examples of the genitive absolute after all. Besides, if the user seeking the genitive absolute stipulates a substantive and a participle, both in the genitive and otherwise in agreement, in either order, within a contextfield of so many words, and without intervening finite verb, how will this pull out the handful of examples where the genitive absolute is "defective," e.g. where it has no

substantive? In other words, the user must do a fair bit of work in the grammars before attempting to write a command file for a particular construction. Without a detailed knowledge of the parameters of the construction, the user will almost certainly write an OPTIONS section that is too restrictive or impossibly broad—the former leaving out many genuine occurrences, and the latter generating enormous output, only a few instances of which are genuine examples of the targetted construction. By the same token, the user must be thoroughly immersed in the complexities of the OPTIONS section of GRAMCORD, or experience enormous frustration at being unable to focus with reasonable precision on the desired construction.

The GRAMCORD package comes with a fairly detailed manual of about 115 pages that explains the terms and concepts necessary to write responsible command files, and that provides a fair number of examples and warnings. For most complex constructions, however, the user must be prepared for output that is several times longer than the list of genuine entries. In other words, it is almost always necessary in such instances to go through the output and cull out the genuine examples. But at least with this tool and a little practice, one may be certain of getting *all* the examples.

GRAMCORD comes with a number of auxiliary programs, the most important of which is GRAMDIR, which enables the user to look up any verse in the NT to see how the GRAMCORD system tags any particular word. The entire package has been adapted to most mainframes, to IBM micros and compatibles, and to a number of other machines. Work is now underway to enable it to run on the MacPlus. In this instance, provided suitable arrangements can be worked out with the publishers of NA<sup>26</sup>, we may be able to provide a version that displays Greek, not transliteration in capital letters.

Certainly the Institute has the NA<sup>26</sup> in Greek font for its own research purposes. Since Carson joined the project in 1979, one of the aims has been to prepare a large volume of syntactical constructions, arranged more or less like a concordance. The University of Chicago Press has agreed to publish *A Syntactical Concordance to the Greek New Testament*. The ms should be ready for the Press early in 1988. Keyed in part to BDF but interacting with most of the major grammars of NT Greek written in the last two centuries, this large volume (approx. 1600 large pages, double column) will overlap as little as possible with the projected volumes of the Friberg project, but will provide exhaustive lists of hundreds of complex syntactical units, complete with a listing of textual variants where those variants affect the construction, and annotated with enough comments to explain the critical decisions that included or excluded particular examples. It must be stressed that the *Concordance* is not raw computer output, with a computer's insensitive handling of

(for instance) the "window" around the target (e.g. automatically printing, say, five words of text on either side of the construction, regardless of whether or not the result can be scanned as comprehensible Greek). Rather, each entry has been examined, and enough text has been included to make the printed line(s) readable. The target elements are made to stand out by being printed in italicized Greek. Where warranted, textual variants follow the entry. Further tags in the margin provide more detailed analysis of the construction, and an asterisk warns the reader that this particular entry receives some mention in the notes that follow each concordance article. In preparation for this work, it was necessary to prepare a multi-lingual font, full-feature editor (including automatic footnoting, word and page wrap, etc.). The editor, called GRAMEDIT, was the first of its kind in a micro environment. Its use is restricted to the NEC APC II; but scholars with access to that machine will find it a powerful aid to multi-font composition and editing.

The GRAMCORD package (not including the editor, which can be purchased separately) can be leased by institutions (where there is a multi-user environment, and some possibility of recovering costs by assessing mild fees) for \$500.00 per annum, which includes unlimited support, free updates as they become available, and several manuals. Private scholars (no multi-user environment) may purchase the package for a one-time fee of \$500.00; updates are extra, and support is more limited. There is also a "GRAMCORD Junior" version available at half the cost. This can retrieve anything that the full version can manage, but because it handles only one construct at a time it takes much longer to retrieve complex syntactical units. Other programs are available, including the GRAMCORD TOOLBOX—a set of programs for column-based sorting, extractions and manipulations of various kinds. Special arrangements are sometimes made with institutions and scholars willing and able to engage in mutually beneficial research.

For full details of the GRAMCORD Project, including lists and prices of available software and application to be placed on a mailing list that notifies interested parties of new developments, write to:

The Director  
The GRAMCORD Institute  
2065 Half Day Road  
Deerfield, IL 60015  
USA

D. A. CARSON  
PAUL A. MILLER