The AGNT Project Report—Q3 2009

As a licensee or friend of AGNT or ANLEX, we would like to update you once a quarter about our continuing work to enhance and perfect these databases and about our plans for the future.

The Project. The Analytical Greek New Testament Project is now in its thirty-third year. Its purpose is to provide aids to translators, students, and scholars of the Greek New Testament.

The initial product was the *Analytical Greek New Testament* (AGNT), which was published by Baker Book House in 1981 and now offered by Trafford Publishing. The second product was the *Analytical Concordance of the Greek New Testament*, published in two volumes (lexical focus; grammatical focus) by Baker Book House in 1991. These two volumes are out of print and are not to be reprinted. Due to advances in computing between 1980 and 1990—to say nothing of those through the present time, there was no need to represent electronically the exact form of the concordances, since their content and more were available through user-initiated interactive searches in the Greek text enabled through several good software packages. The third output of the project was the *Analytical Lexicon of the Greek New Testament* (ANLEX), published in 2000 by Baker Book House and reissued by Trafford in 2005.

The AGNT Project Report—Q3 2008 introduced the team, outlined ongoing tasks, and discussed potential tasks.

The Emdros Linguistic Search Engine. In this issue, Ulrik Sandborg-Petersen, PhD, explains how his program Emdros can assist Bible students in analyzing phrase- and clause-level structure.

The Emdros Linguistic Search Engine

Ulrik Sandborg-Petersen, PhD

For many years, linguistic databases containing analyses of the biblical texts were predominantly concerned with word-level information. For example, the AGNT database—as well as the GRAMCORD database and the Westminster Hebrew Morphology—only contain information relating to each word: its morphological categories, lemma, and other word-level information. In recent years, however, there has been a growing interest within biblical studies circles in going beyond word-level into the realm of language called "syntax."

Syntax deals with how sentences are constructed. Categories such as *clauses* and *phrases* appear in theories concerning syntax. For example, a clause is sometimes

defined as being "a complete thought unit," in which at least a predicate and often a subject is present. The predicate and the subject are usually contained in what linguists call "phrase units." For example, in the clause "Paul sent Timothy to Ephesus" the subject is the phrase "Paul," the predicate is "sent," the object is "Timothy," and the last two words—"to Ephesus"—make up what linguists call a "complement." All four of these units are phrases in a syntactic description of this clause.

Linguistic syntax is often described in terms of "trees." In this context a tree has its "leaves" at the bottom and its "root" at the top, quite the opposite of real trees. In addition, the leaves are usually words, and the root is usually a sentence. Our example sentence may be depicted syntactically as in Figure 1.
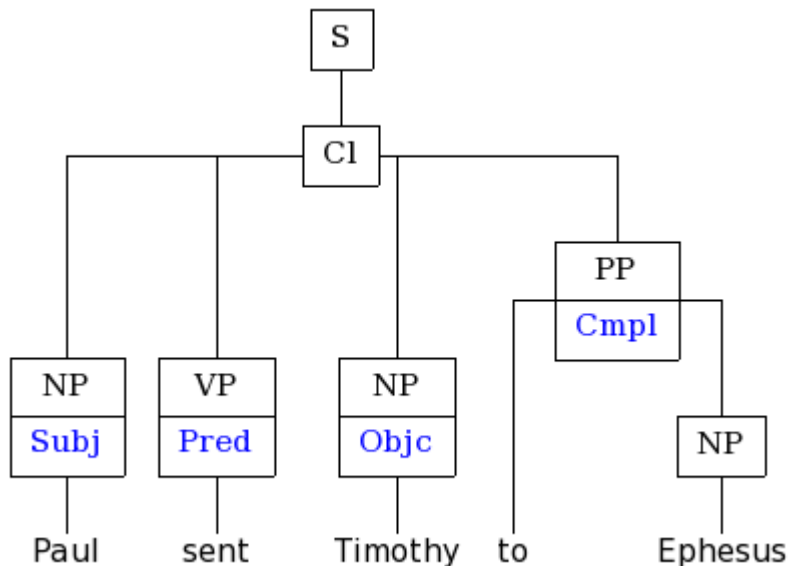

Figure 1: A syntax tree

The advantage of having syntax-level information in addition to the word-level information may not be obvious from the example given. After all, can't we all determine the parts of the sentence in this example—subject, predicate, object, and complement? Yes, but when the underlying language is Greek or Hebrew, and when the examples get only slightly more complicated, the Bible student can be greatly aided in his study of the biblical text by a linguistic analysis that shows diagrammatically (by means of trees) how each sentence is put together.

Issues of exegesis also come up. A syntactic database can—by means of its trees—provide aids to the Bible student in making sense of a difficult sentence or even passage, simply because the relationships between the various parts of the sentence are made explicit, and also because the boundaries of the parts of the sentence (subject, predicate, etc.) are clearly marked.

In addition to the aspect of understanding the biblical texts better, issues of research also come up. For example, not all aspects of Greek or Hebrew grammar are equally well understood or even adequately documented in the standard reference grammars. With syntactic databases at the Bible researcher's disposal, doing research on Greek

or Hebrew grammar becomes a lot easier, especially if the databases are made searchable.

This is where the Emdros linguistic search engine comes in. Emdros is a piece of software that can be embedded in other software, driving linguistic searches of biblical (or any other) texts that have been analyzed linguistically. Emdros is "generic," in that it does not presuppose a certain linguistic theory or way of analyzing. Instead, it leaves it up to the database designer to use the generic facilities that Emdros provides for storing linguistic data. Thus Emdros provides a means for storing any kind of linguistic annotation of text—from morpheme-level to word-level to syntax-level and even to discourse-level.

In addition to merely *storing* linguistic analyses of text, Emdros is first and foremost a *search engine*. Emdros provides a query language thatis intuitive to many after a five-minute explanation. For example, in order to search for two adjacent words, this query might be interesting:

[Word part_of_speech = article]
[Word part_of_speech = noun]

This finds two adjacent words. The first word must have part of speech "article," and the second must have part of speech "noun."

In order to search for words that are embedded inside a phrase, one can simply surround the words with a phrase block, contained in [square brackets]:

[Phrase
        phrase_type = NP

        [Word part_of_speech = article]
        [Word part_of_speech = noun]
]

This would find all phrases whose phrase type was "NP" (noun phrase), and somewhere inside of which we found an article next to a noun.

This can get arbitrarily complex, such that even the most convoluted syntactic expressions can be searched for. The following is slightly more complex than the above, but by no means exhausts the possibilities:

[Clause
        [Phrase function = Subj
                [Word FIRST AND LAST   lemma = "Paul"]
        ]
        [Phrase function = Pred]
        [Phrase function = Objc]
        [Phrase function = Cmpl AND phrase_type = PP
                [Word FIRST part_of_speech = preposition AND lemma = "to"]
        ]
]

This query would, in a properly analyzed Emdros database, find the example sentence with which we began, namely "Paul sent Timothy to Ephesus."

Emdros was developed by me, Ulrik Sandborg-Petersen. Emdros began life in 1999, and is mature software. A number of companies and institutions have licensed Emdros, including Logos Bible Software, the German Bible Society, and Aalborg University. Emdros is in use in academia and research around the world, including by the Werkgroep Informatica in Amsterdam, led by Professor Dr. Eep Talstra, as well as the Groves Center for Advanced Biblical Research, directed by Kirk Lowery, PhD.

A number of syntactically analyzed databases of the biblical texts exist. Some of these are available for licensing by third parties, including the Westminster Hebrew Syntax done by the J. Alan Groves Center for Advanced Biblical Research (http://www.grovescenter.org), directed by Dr. Kirk Lowery. The Groves Center also has a syntactic database of the Greek New Testament in the works. The Groves Center collaborates closely with me on a number of projects, and the Westminster Hebrew Syntax is thus already available in Emdros format. The Groves Center uses the Emdros form of the database as their primary research instrument when analyzing their Hebrew data.

Logos Bible Software has, as was mentioned, also licensed Emdros. Within Libronix, Emdros is used to drive the "syntax search." Similarly, the German Bible Society has licensed Emdros for use within the Stuttgart Electronic Study Bible, in which it drives the "syntax search" of the WIVU syntactic database of the Hebrew Bible.

Emdros is available free of charge under an Open Source license, the General Public License (GPL). The GPL requires that, if you distribute software linked with GPL'ed software, you must distribute your own software under the terms of the GPL, i.e., as Open Source. Commercial companies and institutions that do not wish their software to become Open Source, but who nevertheless wish to license Emdros under a commercial license, are very welcome to contact the author at <ulrikp@emergence.dk>.

Emdros is applicable not only to syntactic databases, but also to word-level databases such as the Analytical Greek New Testament. Since Emdros is a search engine for linguistically analyzed texts, it can as easily be brought to bear on word-level analyses as it can on syntax-level analyses. The strength of Emdros lies in making it easy for the programmer to offer advanced search capabilities, both morphological and syntactic, to the end-user. This presupposes a properly tagged text, of course, such as the Analytical Greek New Testament or the Westminster Hebrew Syntax.

For more information, please visit Emdros's website: http://emdros.org

Ample documentation and examples can be found at the website. The author's company, which handles all licensing of Emdros, has a homepage at: http://emergence.dk

I wish to thank Timothy Friberg and John Hughes for allowing me to make this guest appearance in the AGNT newsletter.

As always, we remain open to developing AGNT and ANLEX in ways that are most useful to the needs of students and readers of God's Word.

Thank you for your continued support of *The AGNT Project*, for faithfully marketing the AGNT and ANLEX databases, and for making these state-of-the-art tools for studying the Greek New Testament available to students, scholars, pastors, and laypeople worldwide.

John Hughes
Agent for *The AGNT Project*
johnhughes@centurytel.net
Phone: 406.862.7289
FAX: 406.862.0917