



Processing Tools for Greek and Other Languages of the Christian Middle East

Bastien Kindt

Research Assistant – Université catholique de Louvain, Belgium

bastien.kindt@uclouvain.be

Abstract

This paper presents some *computer tools* and *linguistic resources* of the GREgORI project. These developments allow automated processing of texts written in the main languages of the Christian Middel East, such as Greek, Arabic, Syriac, Armenian and Georgian. The main goal is to provide scholars with tools (lemmatized indexes and concordances) making corpus-based linguistic information available. It focuses on the questions of text processing, lemmatization, information retrieval, and bitext alignment.

Keywords

Lemmatization; Greek; Syriac; Arabic; Armenian; Georgian; lexical tagging; POS tagging; Unitex; concordances; indexes; finite state transducers; bitext; bilingual alignment; translation memories; Unitex; mkAlign.

INTRODUCTION

The GREgORI project, carried out for many years at the Oriental Institute of the Université catholique de Louvain in Louvain-la-Neuve (Belgium) (https://uclouvain.be/fr/instituts-recherche/incal/ciol/gregori-project.html), brings together linguists, philologists and computer scientists around a common goal: developing *computer tools* and *linguistic resources* allowing automated processing of texts written in Greek [Coulie, 1996; Kindt, 2004; Kindt, 2010] and in the languages of the Christian Middle East, such as Middle Arabic [Tuerlinckx, 2004], Syriac, Armenian and Georgian [Coulie, Kindt and Pataridze, 2013]. From the point of view of computational approach all these languages still remain, especially in the case of the ancient texts, poorly resourced. These *tools* and *resources* make it possible to turn raw texts into tagged corpora enriched with lexical (lexical entry or lemma), morphosyntactical (noun, verb, adjective, etc.), and inflectional (case, gender, number, voice, tense, person, etc.) labels, defining each word-form of processed texts.

This work can lead to both monolingual and multilingual corpora. In the first case, tagged data is used in creating lemmatized concordances and indexes. It is also exploited with corpus

analysis software or data retrieval system. In the second case, the analysis is focused on translation methods. Lexical data of a *source text* is related to lexical data of a *target text* in order to highlight how lexical features from a source language (in relation with a specific cultural area) are spread and received in a target language (linked with another cultural area). In both cases, the main purpose is to provide researchers with linguistic information exclusively drawn from corpus-based evidences.

This paper focuses on the Unitex corpus processor, on the one hand, and on specific outputs such as monolingual or bilingual concordances, indexes and other lexicographical lists provided in the shape of PDF documents, on the other hand. Since two contributions included in the present issue, [Pataridze and Kindt, 2017] (*Text Alignment in Ancient Greek and Georgian: A Case-Study on the First Homily of Gregory of Nazianzus*) and [Van Elverdinghe, 2017] (*Recurrent Pattern Modelling in a Corpus of Armenian Manuscript Colophons*), are founded on the developments presented here, this paper should be regarded as a "general introduction" to these two more specific studies. Given that these contributions concern Armenian and Georgian languages, this paper deals with Greek and Syriac languages.

I. TEXT PROCESSING AND LEMMATIZING

Unitex is an open source corpus processor [Paumier, 2016 and http://unitexgramlab.org/fr]. It supports the Unicode encoding standard and is language independent. This software suite works with specific resources, such as *dictionaries* [Paumier, 2106:43-74] and *grammars* [Paumier, 2106:93-160]. *Dictionaries* provide word-forms labeled with lemmata and grammatical tags (part-of-speech tags, inflectional tags). *Grammars* are user-friendly graphical representations of lexical or syntactic patterns based on the formalism of recursive transition networks.

1.1 Unitex

The following lines illustrate how Unitex allows the user to open, to process, and to explore a Greek corpus presently composed by two hundred and forty-nine letters of Gregory of Nazianzus (329-390) published by [Gallay, 1964-1967] and [Gallay and Jourjon, 1976]. This data includes 45,971 word-forms (11,318 different forms; 4,452 lemmas). This corpus is part of a bigger corpus entitled *Corpus Epistularum Patrum Cappadocum – Corpus épistolographique des Pères cappadociens*, bringing together the letters of Basilius of Caesarea, Gregory of Nyssa, Gregory of Nazianzus, and Firmus of Caesarea, henceforth available on the website of the GREgORI project.

When user opens a corpus, the interface proposes a three-step preprocessing: 1) splitting sentences, 2) normalization of the special words or expressions, and 3) lexical lookup.

Unitex deals with the sentence as a linguistic unit. The texts are split up into sentences based on punctuation. This starting process is followed by the normalization task aimed at improving analysis of some words or expressions. In Greek, this step deals with the *crasis* or *multiword expressions* containing one lexical unit corresponding (out of context) to more than one analysis, but being non ambiguous in the given sequence.

The graphical forms of the *crasis* combine two lexical units written as a single word. The process has to provide an analysis identifying clearly the presence of these two lemmata in the single form. For instance, the forms κἀγὼ "and Ι" and τοὕδαφος "the ground-floor" are processed offering a lexical description including, respectively, καί "and" and ἐγὼ "Ι" and ὁ "the" (form τὸ) and ἔδαφος "ground-floor". After preprocessing, the word-form κἀγὼ in the sentence Πλὴν ὥσπερ ὑμεῖς κύριοι τῆς ὑμετέρας γνώμης, οὕτω κἀγώ τῆς ἐμῆς (Letter 153, 2)

is rewritten {@κἀγώ@,2.K} {καί,καί.I+Part} {ἐγώ,ἐγώ.PRO+Per1s} and τοὕδαφος, in the sentence κόπτω μὲν οἶον τοὕδαφος τοῖς ποσί (Letter 153, 2), is rewritten {@τοὕδαφος@,2.K} {τό,ὁ.DET:Nns:Ans} {ἔδαφος,ἔδαφος.N+Com}. These changes are exclusively applied to the text handled by Unitex and do not modify the original version provided by the published edition. In any case, the concordances are displaying the original Greek texts.

Among the *expressions* containing one ambiguous word without context, we can quote $\dot{\omega}_{\varsigma}$ οἶμαι "in my opinion" or $\ddot{\omega}_{\varsigma}$ φησι "as we say". The word-form $\dot{\omega}_{\varsigma}$ is either the conjunction $\dot{\omega}_{\varsigma}$ "as" or the relatively less frequent preposition $\dot{\omega}_{\varsigma}$ "in, by", synonymous with the more common εἰς, but exclusively used with a name of person or place such as in the sentence ἀφίκετο $\dot{\omega}_{\varsigma}$ Περδίκκαν καὶ ἐς τὴν Χαλκιδικήν "he arrived at Perdiccas in Chalkidiki" [Liddell-Scott, 1958, p. 2039]. Followed by οἶμαι or φησι the word-form $\dot{\omega}_{\varsigma}$ is unequivocally the conjunction. For these two expressions, the sentences are rewritten, after preprocessing, as κεκλήκατε ἡμᾶς ἐπὶ τὴν μητρόπολιν, { $\dot{\omega}_{\varsigma}$, $\dot{\omega}_{\varsigma}$ ($\ddot{\omega}_{\varsigma}$).I+Conj} {οἷμαι,οἵομαι ($-\omega$).V}, περὶ ἐπισκόπου τι βουλευσόμενοι (Letter 43, 1) and ἀλλὰ νῦν μέν, { $\ddot{\omega}_{\varsigma}$, $\dot{\omega}_{\varsigma}$ ($\ddot{\omega}_{\varsigma}$).I+Conj} {φησι,φημί.V} Πίνδαρος (204, 5, 147, 23). Once again, the change only concerns the text handled by Unitex.

The third step of the preprocessing involves the lexicon lookup. The GREgORI project develops its own dictionaries for each processed language. The dictionaries are Unicode text files listing the word-forms. The Greek lexicon records 441,464 word-forms connected with 67,347 lemmata and, to be used with Unitex, is converted into DELAF format [Paumier, 2016:43-46]. Each word-form is provided with a one-line lexical entry using a precise syntax: inflected form, lemma corresponding to the form, part-of-speech tag of the lemma (noun, verb, adjective, adverb, pronoun, etc.) and inflectional tags describing the inflected-form (case, gender, number, voice, tense, mood, person, etc.). For instance: τό,ὁ.DET:Nns:Ans; ἔδαφος,ἔδαφος.N+Com:Nns:Vns:Ans, etc. The DELAF formalism is also used for the other languages, such as Armenian npnpp,npnp.N+Com, Syriac κίω, κίω. NOUN or Georgian ξεδημος, ξεδημος. N+Com, these three lemmata designating "the son".

It is important to precise that the lexical lookup does not take into account the context in which the words appear. Therefore, when a word corresponds to more than one lemma, all the possibilities are memorized by the process, such as, for example, the inflected form $\phi \dot{\phi} \sigma \epsilon$, corresponding to both the verb $\phi \dot{\phi} \omega$ "engender" and the noun $\phi \dot{\phi} \sigma \zeta$ "nature". This is a case of lexical ambiguity (see sections 1.3 and 1.5, below).

1.2 Information retrieval

After the preprocessing, the user can search the text for words or expressions by using the "Locate Pattern" menu [Paumier, 2016:84-89]. The results are displayed in the shape of "Key Word In Context" concordances. Queries may concern inflected forms, lemmata, part-of-speech tags but also the combination of these arguments, as shown in table 1 below.

Locate Pattern	Result
εἴχομεν	The concordance of the word-from εἴχομεν
činos.	The concordance of all the inflected form of the verb
<ωχώ>	έχω
<v></v>	The concordance of all the inflected form of lemmata
< v >	tagged as verb
L Non cours	The concordance of all the lemmata tagged as negation
<i+neg><ἔχω></i+neg>	followed by an inflected form of the verb ἔχω

Table 1. Samples of queries.

Fig. 1 below gives a part of the concordance corresponding to the query <I+Neg><ἔχω>. Note once again that, at this stage of the process, before lexical disambiguation, queries on both <φύω> and <φύσις> provide concordances including the word-form φύσει.

[GrNa-29-9-26-25]	εἰ δόξειε μόνος τῶν πάντων Καισάριος	μὴ ἐσχηκέναι	φίλους, ό δὴ καὶ πολλοὺς ἔχειν
[GrNa-176-1-126-18]	έγκαλέσομεν, οὔτε ἀργίαν, οὔτε τὸ	μὴ ἔχειν	ő τι γράψεις·{S} οὐ πρὸς σοῦ τὸ ταῦτα
[GrNa-202-14-92-11]	έκεῖνον τὸν ἄνωθεν ἥκοντα τὸν νοῦν	μὴ ἔχειν,	ἀλλὰ τὴν θεότητα τοῦ Μονογενοῦς τὴν τοῦ
[GrNa-249-19-181-14]	ὑπεροχὴ τὰ χαυνότερα ἤθη διαφυσήσῃ,	οὐκ εἶχον	ὅπως ἐμαυτὸν ἀτρεμεῖν νουθετήσω,
[GrNa-202-20-94-7]	τοῦ αὐτοῦ πράγματος ἀληθεῖς εἶναι φύσιν	οὐκ ἔχει.{S}	Πῶς οὖν ὑπέμεινέ σου ἡ μεγαλοφυὴς καὶ
[GrNa-170-1-123-5]	γραμμάτων ὑμῖν ὁμιλεῖν, ἐπειδή γε ἄλλως	οὐκ ἔχομεν,	οὕτως οἰκονομήσαντος {τοῦ,ὁ.DET}
[GrNa-79-7-70-13]	καὶ τὴν πνευματικὴν ἐπιστασίαν	οὐκ ἔχοντος;{S}	Ἐκεῖνο δέ σοι καὶ πρὸ πάντων δῆλον, ὅτι
[GrNa-96-1-79-22]	Ύπατίω Πολὺν ἐζημιώμεθα χρόνον,	οὐκ ἐχούσης	τὸν πρῶτον ἐν ἀνδράσι τῆς πρώτης ἐν
[GrNa-82-2-72-17]	μαχομένοις αὐτοῖς ἥ τι πλέον, τὸ ἄθλον	οὐκ ἔχουσι	τοῦ παλαίσματος.{S} Τῷ αὐτῷ Ἐπαινῶ ὅτι
[GrNa-31-7-28-13]	λέγω·{S} τὸ δὲ ἐζημιώθην·{S}	οὐκ ἔχω	οἵου σὺ χρήζεις, τὴν Ἰλιάδα.{S} Μὴ γάρ μοι

Figure 1. Concordance of the lemmata tagged as negation and followed by an inflected form of the verb ἔχω.

1.3 Text Automaton

Unitex provides a visual representation of each sentence of the processed text, called "automaton" [Paumier, 2106:161-198], as shown on fig. 2. It begins with an arrow and ends with a square inscribed in a circle. Each word form of the sentence appears in a box. Links between boxes symbolize the continuity of the sentence. Each word-form is accompanied with the corresponding lemma and part-of-speech tag, information coming from the lexical lookup.

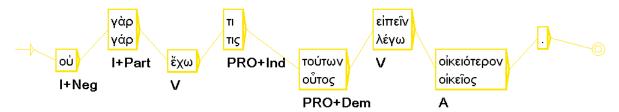


Figure 2. Automaton of the sentence οὐ γὰρ ἔχω τι τούτων εἰπεῖν οἰκειότερον. (Letter 1, 1).

When a word corresponds to more than one lemma, two or more boxes appear on the same horizontal axis, making different interpretations of words explicit. The automaton in fig. 3 shows the lexical ambiguity of the word-form π έτρας, inflected form of Πέτρα "Petra" (the city of Petra), or of π έτρα "stone".

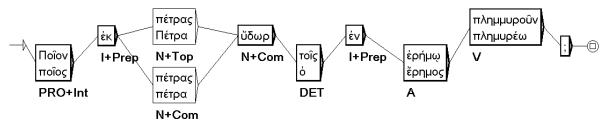


Figure 3. Automaton of the sentence Ποῖον ἐκ πέτρας ὕδωρ τοῖς ἐν ἐρήμω πλημμυροῦν; (Letter 44, 2).

The processed text is fully lemmatized, without lexical ambiguities, when the automaton shows only one box for each word of the sentences of the text, i.e. only one interpretation. At this stage, it is possible to generate valid statistics and to edit lemmatized concordances or other outputs. Lexical ambiguities are processed automatically or manually. In the second case, the user can use the lemmatization interface of Unitex, as shown in section 1.5 below.

1.4 Information retrieval by grammars

Grammars are powerful tools able to describe sequences of words. The query <I+Neg><ἔχω> used above can be graphically represented in the shape of a grammar, as shown by fig. 4. The graphical interface of Unitex allows users to draw grammars by themselves.



Figure 4. Grammar corresponding to the query <I+Neg><ἔχω>.

By applying this grammar to the text with Unitex, the user will be provided, once again, with the concordance of fig. 1. However, the grammars allow to search more complex patterns corresponding to a wide range of possible searches.

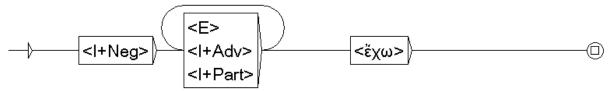


Figure 5. Grammar predicting the optional presence of an adverb or a particle between the arguments <I+Neg> and <ĕ χ ω >.

The grammar of fig. 5 above will furnish the concordance of all the sequences constituted by a negation followed by the inflected forms of the verb $\xi \chi \omega$, just like the previous one, but predicts, between these two arguments, the optional presence (the sign $\langle E \rangle$ allows to recognize an empty sequence) of one or several adverbs or particles.

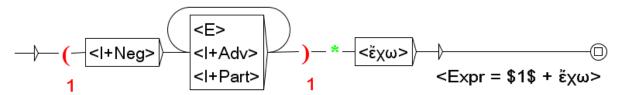


Figure 6. Grammar providing outputs and alphabetical sorting of inflected forms of the last verbal argument.

The grammars can also provide the outputs in the shape of the tags included in the results (fig. 6 above). These tags follow or surround the sequences displayed by the concordance. This grammar provides the same concordance as above, but inflected forms of $\xi\chi\omega$ are classified according to the alphabetical order (due to the use of the asterisk before $\langle\xi\chi\omega\rangle$). Here, the use of variables allows to repeat in the output the part of the sequence inscribed between the brackets, as shown on the concordance provided below (fig. 7).

[GrNa-249-19-181-14]	χαυνότερα ἤθη διαφυσήσῃ,	οὐκ εἶχον <Εxpr=οὐκ+ἔχω>	ὅπως ἐμαυτὸν ἀτρεμεῖν
[GrNa-29-9-26-25]	τῶν πάντων Καισάριος	μὴ ἐσχηκέναι <Εχρι=μὴ+ἔχω>	φίλους, ὁ δὴ καὶ πολλοὺς
[GrNa-81-1-72-4]	Μηδαμῶς, ὧ θαυμάσιε,	μὴ οὔτως ἔχε <Εxpr=μὴ οὔτως+ἔχω>.{S}	Τῶν μὲν γὰρ ἀκούσιος ἡ
[GrNa-202-20-94-7]	ἀληθεῖς εἶναι φύσιν	οὐκ ἔχει <Εxpr=οὐκ+ἔχω>.{S}	Πῶς οὖν ὑπέμεινέ σου ἡ
[GrNa-162-3-117-13]	Νῦν δὲ δέδοικα μὴ	ούχ οὕτως ἔχει <Εxpr=ούχ οὕτως+ἔχω>-{S}	μάλιστα μὲν γὰρ οὐδὲ μεμέ
[GrNa-176-1-126-18]	οὔτε ἀργίαν, οὔτε τὸ	μὴ ἔχειν <Εχρι=μὴ+ἔχω>	ő τι γράψεις·{S} οὐ πρὸς
[GrNa-202-14-92-11]	ἄνωθεν ἥκοντα τὸν voûv	μὴ ἔχειν <Εχρι=μὴ+ἔχω>,	άλλὰ τὴν θεότητα τοῦ
[GrNa-170-1-123-5]	όμιλεῖν, ἐπειδή γε ἄλλως	οὐκ ἔχομεν <Εχρι=οὐκ+ἔχω>,	οὕτως οἰκονομήσαντος
[GrNa-215-4-155-7]	ὑψηλότερον.{S} Ἡμεῖς δὲ	ούχ οὔτως ἔχομεν <Εχρι=οὐχ οὔτως+ἔχω>,	ἀλλ´ ὅτι τοῦ ἑνὸς διημάρτο
[GrNa-79-7-70-13]	πνευματικὴν ἐπιστασίαν	οὐκ ἔχοντος <Εxpr=οὐκ+ἔχω>;{S}	Έκεῖνο δέ σοι καὶ πρὸ

Figure 7. Concordance with outputs and inflected forms of ἔχω in alphabetical order.

In fact, the preprocessing described above (section 1.1) works with this kind of grammar to analyze crasis (fig. 8 and 9) and expressions (fig. 10). In that case, the sequences are fully replaced by the content of the grammar.

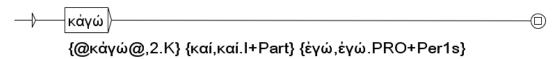


Figure 8. Grammar for the crasis κἀγώ.



Figure 9. Grammar for the crasis τοὕδαφος.

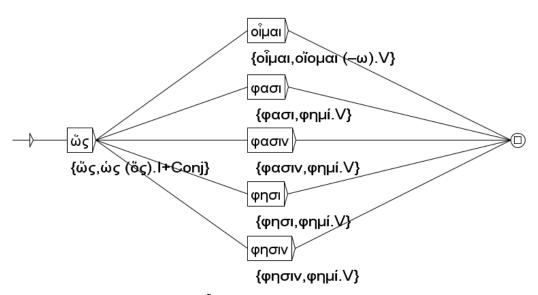


Figure 10. Grammar for the expressions ὡς οἶμαι and ὡς φησι.

Since Unitex uses the Unicode encoding and is language independent, as written above, this kind of grammars can be adapted to analyse similar linguistic facts in Armenian or Georgian. As regards Armenian, a case study is provided by [Van Elverdinghe, 2017], in this issue.

1.5 Lemmatization

Unitex provides a lemmatization interface (fig. 11) allowing to resolve lexical ambiguities manually. A special window displays the concordance and the text automaton of sentences containing lexical ambiguities.

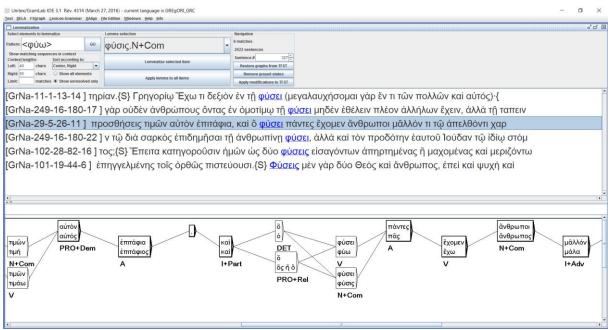


Figure 11. Lemmatization interface of Unitex.

By clicking on superfluous boxes – for instance, boxes with lemma "ὁ.DET" and lemma "φύω.V", the user can remove inadequate interpretations and solve ambiguities in order to obtain a fully lemmatized and disambiguated text, as shown in fig. 12.

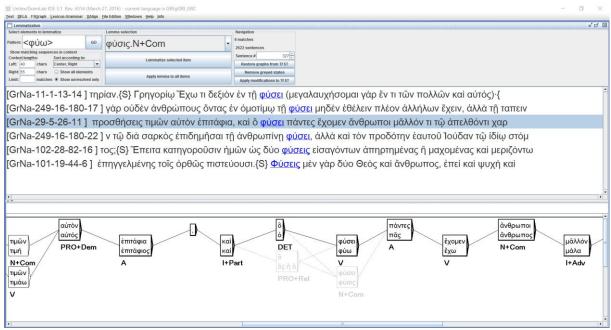


Figure 12. Lemmatization interface with elimination of useless boxes.

All linguistic resources, dictionaries and grammars, can be created, updated and corrected by the users, without any special knowledge of computer programming. The linguistic data is not encapsulated in a "black box" and remains accessible and readable.

At the end of the process, all lemmatized and disambiguated data are exported and gathered in a relational database system. Then, the publishing tools (see section III) allow to edit the special outputs in PDF format, such as concordances, indexes and other lexicographical lists.

II. TEXT ALIGMENT

The GREgORI project processes texts in Greek, Arabic, Syriac, Armenian and Georgian using appropriate resources for each language. In this context, we are interested in comparing the lexical data of a *source text* with its ancient or modern translations, the *target text*, in order to provide scholars with corpus-based information about translation methods. This task requires alignment software such as mkAlign [Fleury, 2012 and http://www.tal.univ-paris3.fr/mkAlign/]. The mkAlign software allows users to pair words or expressions of a source text with words or expressions of a target text.

The mkAlign interface is divided into two columns. The user loads the source text in the left column and the target text in the right one. The first alignment is automatically provided on the basis of the punctuation or textual references of the edition (chapters, paragraphs, etc.). The sample provided here concerns the Greek text of the Homily 13 by Gregory of Nazianzus, aligned with its Syriac (S2) version [Schmidt, 2002]. Then, the user can move the words or expressions to align the relevant translation units. Word by word alignment is not always possible, as shown on fig. 13, below, where the Greek form αἰρόμεναι is aligned with three Syriac words. Aligned data are afterwards exported as bitext files in the Translation Memory eXchange format (.tmx) and gathered in the relational database system. The bitext alignment is currently done manually. In the near future, the bilingual data will be used as translation memories in order to automatize the alignment for recurring translation units.

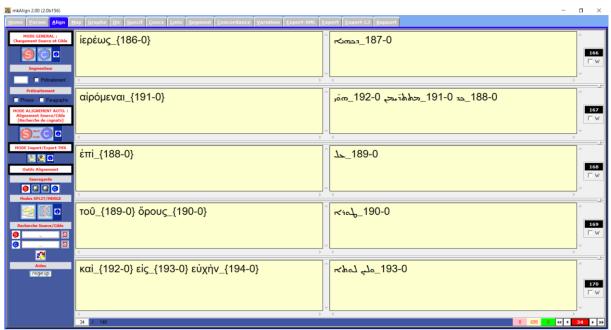


Figure 13. The mkAlign interface.

III. CONCORDANCES, INDEXES AND LEXICOGRAPHICAL LISTS

Specific tools are developed by the Centre de traitement automatique du langage (Cental) at the UCL (https://uclouvain.be/fr/instituts-recherche/ilc/cental) in order to edit concordances, indexes and other types of lexicographical lists. The purpose of these developments is to extract

the lexical data out of the databases and inject them in the frameworks corresponding to the output required by the user.

Fig. 14 shows the lemmatized concordance of the verbs included in the letters of Gregory of Nazianzus. In this PDF document, lemmata are displayed in green (accompanied with a frequency rate), POS tags in red and word-forms in blue, with downstream and upstream contexts. The references to the edition appear on the left of the downstream context. The bookmarks list lemmata and allow browsing the document.

Fig. 15 gives a sample of the alphabetical list of these verbs and Fig. 16 lists these verbs according to their frequency in the corpus.

Fig. 17 offers a sample of a bilingual Greek-Syriac concordance of Homily 13 by Gregory of Nazianzus. This document displays all the lexical data both for the source text and for the target text, including roots (in red) for Syriac. Such a bilingual concordance is the so called "scholar concordance". Fig. 18 provides the lexical data (lemmata and POS tags) only for the source text. Such a bilingual concordance is a so-called "light concordance", being more easily readable for the user.

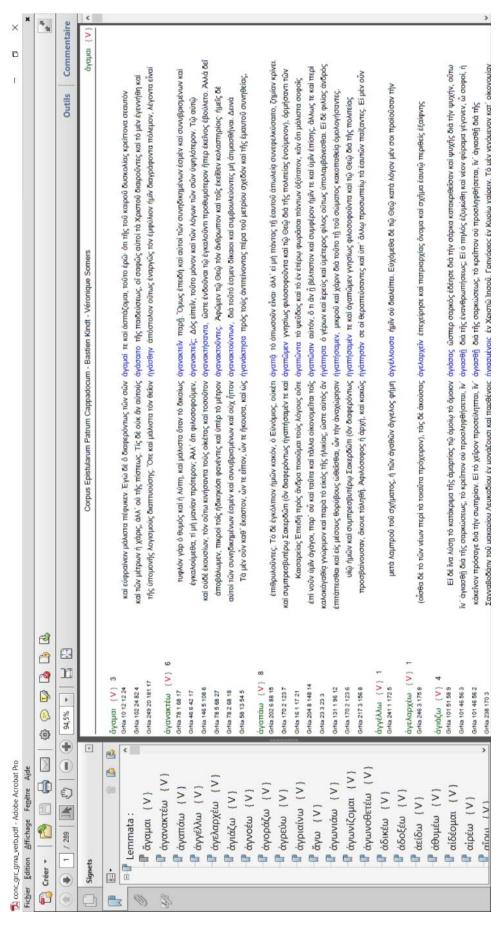


Fig. 14. Lemmatized concordance of the verbs in the corpus of the letters by Gregory of Nazianzus

φησιατι (γ) 3 φληθείων (γ) 1 φισιατρόνουμι (γ) 1 φισιατρόνουμι (γ) 2 φισιατρόνουμι (γ) 1 φληθείων (γ) 2 φισιατρία (γ) 3 φισιατρία (γ) 3 φισιατρία (γ) 3 φισιατρία (γ) 3 φισιατρία (γ) 1 φισιατρία (γ) 2 φισιατρία (γ) 1 φισιατρία (γ) 1 φισιατρία (γ) 2 φισιατρία (γ) 3 φισιατρία (γ) 1 φισιατρία (γ) 2 φισιατρία (γ) 2 φισιατρία (γ) 2 φισιατρία (γ) 3 φισιατρία (γ) 4 φισιατρία (γ) 5 φισιατρία (γ) 5 φισιατρία (γ) 6 φισιατρία (γ) 7 6			Corpus Epistularum Patrum C	appadoc	Corpus Epistularum Patrum Cappadocum - Bastien Kindt - Véronique Somers			
τερο (Ψ) 6 αλείσουμα (Ψ) 1 αναθουμά (Ψ) 1 αναθουμά (Ψ) λω (Ψ) 1 αλληγορία (Ψ) 1 άναθουμα (Ψ) 1 άναθουμα (Ψ) λω (Ψ) 1 αλληγορία (Ψ) 1 άναθομβάνω (Ψ) 1 άναθομβάνω (Ψ) 1 άναθομβάνω (Ψ) το (Ψ) 1 αλληγορία (Ψ) 1 άναθομβάνω (Ψ) 2 άναθομα (Ψ) 3 άναθομα (Ψ) 3 άναθομα (Ψ) το (Ψ) 1 αμειβω (Ψ) 2 άναθεμα (Ψ) 3 άναθεμα (Ψ) 3 άναθεμα (Ψ) 3 άναθεμα (Ψ) λω (Ψ) 2 άμαθο (Ψ) 2 άναθεμα (Ψ) 3 άναθεμα (Ψ) 3 άναθεμα (Ψ) 3 άναθεμα (Ψ) λω (Ψ) 2 άμαθο (Ψ) 1 άναθεμα (Ψ) 1 άναθεμα (Ψ) 1 άναθεμα (Ψ) 1 άναθεμα (Ψ) λω (Ψ) 2 άμαθο (Ψ) 1 άναθεμα (Ψ) 1 άναθεμα (Ψ) 1 άναθεμα (Ψ) 1 άναθεμα (Ψ) λω (Ψ) 2 άναθεμα (Ψ) 1 άναθεμα (Ψ) 1 άναθεμα (Ψ) 2 άναθεμα (Ψ) 2 άναθεμα (Ψ) λω (Ψ) 3 άναθεμα (Ψ) 1 άναθεμα (Ψ) 2 άναθεβλα (Ψ) 2 άναθεβλα (Ψ) 3 άναθεβλα (Ψ) 4 άνασταθεμα (Ψ)	άγαμαι (V)	m	άληθεύω {V}	-	άνακεράννυμι {V}	-	άνεγείρω {V}	-
δια (γ) β άλτοπορια (γ) 2 άναστατια (γ) 3 άνερατάνα (γ) λω (γ) 1 άλληπορρια (γ) 1 άναστατα (γ) 3 άνερατάνα (γ) δω (γ) 1 άλληπορρια (γ) 1 άναστατα (γ) 5 άνερατάνα (γ) δω (γ) 1 άναστατα (γ) 1 άναστατα (γ) 1 άναστατα (γ) δω (γ) 1 άμαν (βερίζω) (γ) 7 άναστα (γ) 3 άντημα (γ) δω (γ) 1 άναστα (γ) 7 άναστα (γ) 3 άντημα (γ) δω (γ) 2 άμαν (βερίζω) (γ) 7 άναστα (γ) 3 άντημα (γ) δω (γ) 2 άμαν (βερίζω) (γ) 7 άναστα (γ) 3 άντημα (γ) δω (γ) 2 άμαν (βερίζω) (γ) 1 άναστα (γ) 3 άντημα (γ) δω (γ) 1 άναστα (γ) 1 άναστα (γ) 1 άναστα (γ) δω (γ) 1 άναστα (γ) 2 άναστα (γ) 1 άναστα (γ) δω (γ) 1 άναστα (γ) 2 άναστα (γ) 4 άναστα (γ) δω (γ) 2 άναστα (γ) 2 άναστα (γ) 4 άναστα (γ) δω (γ) 2 άναστα (γ) 2 άναστα (γ) 4 άναστα (γ) δω (γ) 2 άναστα (γ) 2 άναστα (γ) 3 άναστα (γ) </td <td>άγανακτέω {V}</td> <td>9</td> <td>άλιεύω {V}</td> <td>-</td> <td>άνακομίζω {ν}</td> <td>-</td> <td>άνέλκω {V}</td> <td>2</td>	άγανακτέω {V}	9	άλιεύω {V}	-	άνακομίζω {ν}	-	άνέλκω {V}	2
λω (∀) 1 άλληγορέω (∀) 1 άναναμένω (∀) 5 άνεξω (∀) 1 άναληγορέω (∀) 1 άναληγορέω (∀) 1 άναληγορέω (∀) 2 άναξω (∀) 2 άναξω (∀) 3 άναξω (∀) 3 άναξω (∀) 4 άναξω (∀) 2 άμελδω (∀) 2 άμελδω (∀) 2 άμελδω (∀) 3 άναμηνηνω (∀) 3 άναμην (∀) 4 άναμηνηνω (∀) 3 άναμην (∀) 4 άναμηνηνω (∀) 4 άναμηνηνω (∀) 4 άναμηνηνω (∀) 5 άναμηνηνω (∀) 1 άναμηνηνω (∀) 2 άναμηνην (∀) 2	άγαπάω {V}	00	άλίσκομαι {V}	2	άνακόπτω {V}	3	άνερωτάω {V}	-
(V) 1 αλλοτριώα (V) 1 αναλοιμώαν (V) 1 αναλοιμών (V) 1 αναλοιμών (V) 1 αναλοιμών (V) 1 αναλοιμών (V) 2 αναμένω (V) 2 αναμένω (V) 2 αναμένω (V) 3 αναμένω (V) 3 αναμένω (V) 3 αναμένω (V) 3 αναμένω (V) 4 αναμένω (V) 1 αναμένω (V) 3 αναμένω (V) 4 αναπέρω (V) 1 αναμένω (V) 1 αναμένω (V) 1 αναπέρω (V) 1 αναπέρω (V) 1 αναπέρω (V) 1 αναπέρω (V) 2 αναπέρω (V) 2 αναπέρω (V) 3 αναπέρω (V) 3 αναπέρω (V) 3 αναπέρω (V) 3 αναπέρω (V) 4 αναπέρω (V) 1 αναπέρω (V) 2 αναπέρω (V) 1 αναπέρω (V) 1 αναπέρω (V) 1 αναπέρω (V) 2 αναπέρω (V) 1 αναπέρω (V) 2 αναπέρω (V) 1 αναπέρω (V) 2 αναπέρω (V) 2 αναπέρω (V) 3 αναπέρω (V) 1 αναπέρω (V) 3 αναπέρω (V) 1 αναπέρω (V) 2 αναπέρω (V) 3 αναπέρω (V	άγγέλλω {V}	-	άλληγορέω {V}	-	άνακύπτω {V}	-	άνευρίσκω {V}	3
ω (γ) 4 φιαρτάνω (γ) 7 φυαλίανω (γ) 3 φυήνα (γ) δώα (γ) 1 φιαρτάνω (γ) 5 φυαρίαω (γ) 3 φυήνα (γ) δώα (γ) 1 φιτώρω (γ) 5 φυαρίαω (γ) 3 φυήνα (γ) λω (γ) 2 φιτώρω (γ) 7 φυαρίανημ (γ) 1 φυσίαμη (γ) (γ) 2 φιτώρω (γ) 1 φυσίαμη (γ) 3 φυσίαμη (γ) (γ) 2 φιτώρω (γ) 1 φυσίαμη (γ) 1 φυσίαμη (γ) (γ) 14 φιτφράρτω (γ) 1 φυστάρω (γ) 1 φυσίαμη (γ) δροια (γ) 14 φιτφράρτω (γ) 1 φυστάρω (γ) 1 φυσίαμη (γ) ω (γ) 14 φιτφράρτω (γ) 1 φυστήρω (γ) 1 φυστήρω (γ) ω (γ) 14 φυσράρτω (γ) 1 φυστήρω (γ) 1 φυστήρω (γ) ω (γ) 14 φυσράρτω (γ) 1 φυστήρω (γ) 1 φυστήρω (γ) ω (γ) 15 φυσράρμω (γ) 1 φυστήρω (γ) 1 φυστήρω (γ) ω (γ) 2 φυσράρμω (γ) 1 φυστήρω (γ) 1 φυστήρω (γ) ω (γ) 2 φυσράρμω (γ) 1 φυστήρω (γ) 1 φυστήρω (γ) ω (γ) 2 φυσράρμω (γ) 2 φυστήρω (γ) 2 φυστήρω (γ)	άγελαρχέω {V}	-	άλλοτριόω {V}	-	άναλαμβάνω {V}	5	άνέχω {V}	11
ξω (Υ) 19 φμάω (θερίζω) (Υ) 1 φασβώω (Υ) 3 φασβώω (Υ) ξω (Υ) 1 φμέρω (Ψ) 5 φασμένω (Υ) 6 φαθέω (Υ) λοω (Υ) 2 φμόνω (Υ) 1 φασμένω (Υ) 3 φασβημα (Υ) ξω (Υ) 1 φμόνω (Υ) 2 φασβώνω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) ξωρα (Υ) 2 φμόνω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) ξωρα (Υ) 14 φαρβήτεω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) μα (Υ) 34 φαρβάλω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) μα (Υ) 11 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) μα (Υ) 11 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) 2 φασμένω (Υ) μα (Υ) 11 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) μα (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) μα (Υ) 2 φασβάλω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) 1 φασμένω (Υ) (Υ) 2 φασβάρμα (Υ) 2 φασμένω (Υ) 1 φασμένω (Υ)	άγιάζω {Λ}	4	άμαρτάνω {V}	7	άναλίσκω {V}	-	άνηβάσκω {V}	-
δέφ (γ) 1 άμτεβω (γ) 5 άνσμένω (γ) 6 άνθέω (γ) λω (γ) 2 άμτοβω (γ) 7 άνσμήνητα (γ) 1 άνσημη (γ) ξίνη 2 άμτοβω (γ) 1 άνσμηνησια (γ) 1 άνσημη (γ) (γ) 2 άμτοβω (γ) 1 άνσμηνησια (γ) 1 άνσημη (γ) (γ) 2 άμτοβολω (γ) 1 άνσπεξω (γ) 1 άνστραξω (γ) δορτεία (γ) 1 άνσηκολω (γ) 3 άνσπεξω (γ) 1 άνσημη (γ) ω (γ) 34 άνσηβολω (γ) 4 άνσπλησω (γ) 1 άνσημοζω (γ) ω (γ) 1 άνσηκολω (γ) 4 άνσπλησω (γ) 1 άνσηκολω (γ) ω (γ) 1 άνσηκολω (γ) 2 άνσηκομ (γ) 4 άνσπροξω (γ) μαι (γ) 2 άνσηξω (γ) 4 άνσπροξω (γ) μαι (γ) 2 άνσηξω (γ) 1 άνσπροξω (γ) μαι (γ) 2 άνσηξω (γ) 1 άνσπροξω (γ) μαι (γ) 2 άνσηξω (γ) 1 άνσπροξω (γ) 1 άνσπροξω (γ) μαι (γ) 2 άνσηξω (γ) 1 άνσπροξω (γ) 1 άνσπροξω (γ) μαι (γ) 3 άνσπροξω (γ) 1 άνσπροξω (γ) 1 άνσπροξω (γ) μαι (γ) 3 άνσπροξω (γ) 3	άγνοέω {V}	19	άμάω (θερίζω) {V}	-	άναλύω {V}	3	άνήκω {V}	-
λω (γ) 2 άμελθω (γ) 7 άνσιμιγήνωμι (γ) 1 άνσιμι (γ) (ψ) 29 άμουω (γ) 1 άνσιμη (γ) 3 άνίημι (γ) ζομοι (γ) 2 άμονοβρίτων (γ) 1 άνσιστημι (γ) 1 άνσιστημι (γ) ζομοι (γ) 1 άνσιβρίτων (γ) 3 άνσιτεβω (γ) 7 άνσιβρίτω (γ) ω (γ) 34 άνσιβρίτων (γ) 3 άνσιτεβω (γ) 7 άνσιβρίτω (γ) ω (γ) 11 άνσιβρίτων (γ) 4 άνσιβρίτω (γ) 7 άνσιβρίτω (γ) ω (γ) 11 άνσιβρίτω (γ) 4 άνσιπεβω (γ) 7 άνσιβρίτω (γ) ω (γ) 11 άνσιβρίτω (γ) 1 άνσιπερω (γ) 4 άντερφρομ (γ) ψα (γ) 1 άνσιβρίτω (γ) 1 άνσιπερω (γ) 1 άνσιπερω (γ) ψα (γ) 2 άνσιπερω (γ) 1 άνσιπερω (γ) 1 άντερφρομ (γ) ψα (γ) 2 άνσιπερω (γ) 1 άνσιπερω (γ)	άγοράζω {Λ}	-	άμείβω {V}	2	ávatrévu {V}	9	άνθέω {V}	-
(V) 1 φινημονέω (V) 1 φινημονέω (V) 3 φινήμι (V) (V) 29 φινόνω (V) 2 φινοποίο (-φω) (V) 1 φιναστημι (V) (Φ) 1 φινόπο (V) 1 φιναστήμι (V) 1 φιναστήμι (V) (Φ) 14 φιναθράνω (V) 1 φιναπόμο (V) 1 φιναστήμα (V) αθετέω (V) 34 φιναβράνω (V) 1 φιναστήμα (V) 1 φιναστήμα (V) αθ (V) 11 φιναγγάζω (V) 4 φιναπόμο (V) 4 φιναστήμα (V) εια (V) 11 φιναγγάζω (V) 2 φιναπόμα (V) 4 φιναστήμα (V) εια (V) 11 φιναγγάζω (V) 1 φιναστάμα (V) 4 φιναστήμα (V) εια (V) 1 φιναστάμα (V) 1 φιναστάμα (V) 4 φιναστάμα (V) εία (V) 2 φιναθήμα (V) 1 φιναστάμα (V) 1 φιναστάμα (V) εία (V) 2 φιναθήμα (V) 1 φιναστάμα (V) 1 φιναστάμα (V) εία (V) 2 φιναθήμα (V) 3 φιναστάμα (V) 1 φιναστάμα (V) 1 φιναστάμα (V) εία (V) 3 φιναστάμα (V) 1 φιναστάμα (V) 1 φιναστάμα (V) 1 φιναστάμα (V) εία (V) 15 φινασθόμτα (V) 1 φιναστάμα (V) 1 φιναστάμα (V)	άγρεύω {V}	2	άμελέω {V}	7	άναμίγνυμι {\\}	-	άνιάω {\\V}}	7
(γ) 29 άμόνω (γ) 2 άνασκόραι (-όω) (γ) 1 άνίσχω (γ) cigu (γ) 14 άμφεδνυμι (γ) 1 άνασταίω (γ) 1 άνασταίω (γ) ω (γ) 34 άναβαίνω (γ) 1 άνασταίω (γ) 1 άνασμάσου (γ) ω (γ) 34 άναβαίνω (γ) 4 άνασταίω (γ) 1 άνασμάσου (γ) εία (γ) 11 άναγκάζω (γ) 4 άνασταίω (γ) 4 άνασταίρω (γ) εία (γ) 11 άναγκάζω (γ) 1 άνασταίρω (γ) 4 άνασταίρω (γ) εία (γ) 11 άναγκάζω (γ) 1 άνασταίρω (γ) 4 άνασταξύμα (γ) μαι (γ) 1 άνασταίνω (γ) 1 άνασταξύμα (γ) 1 άνασταξύμα (γ) εί (γ) 5 άναδιξόσκω (γ) 1 άνασταξύμα (γ) 1 άνασταίμα (γ) εί (γ) 5 άνασταίνα (γ) 1 άνασταίμα (γ) 1 άνασταίμα (γ) εί (γ) 5 άνασταίμα (γ) 1 <	άγριαίνω { V }	-	άμνημονέω {V}	-	άναμιμνήσκω {V}	3	άνίημι {V}	5
dάω (∀) 2 diquévolut (∀) 1 docacióu (∀) 1 docacióu (∀) 1 docacióu (∀) βόρια (∀) 14 diqueβητέω (∀) 1 docaraciúa (∀) 2 diqueghytém (∀) 1 docaraciúa (∀) αθ (∀) 34 docacióa (∀) 1 docaracióa (∀) 1 docaracióa (∀) 1 docaracióa (∀) α (∀) 11 docarque (∀) 1 docarcióa (∀) 4 docarcióa (∀) 4 docarcióa (∀) δ (∀) 1 docarcióa (∀) 1 docarcióa (∀) 4 docarcióa (∀) 4 docarcióa (∀) ξ (∀) 27 docarcióa (∀) 1 docarcióa (∀) 4 docarcióa (∀) 4 docarcióa (∀) ξ (∀) 5 docacióa (∀) 1 docarcióa (∀) 4 docarcióa (∀) 4 docarcióa (∀) ξ (∀) 5 docacióa (∀) 1 docarcióa (∀) 4 docarcióa (∀) 4 docarcióa (∀) ξ (∀) 5 docacióa (∀) 1 docarcióa (∀) 1 docarcióa (∀) 1 docarcióa (∀) ξ (∀) 5 docacióa (∀) 1 docarcióa (∀) 1 docarcióa (∀) 1 docarcióa (∀) ξ (∀) 2 docarcióa (∀) 1 docarcióa (∀) 1 docarcióa (∀) 1 docarcióa (∀) δ docarcióa (∀) 2 docarcióa (∀) 1 docarcióa (∀) <t< td=""><td>ãyw {V}</td><td>53</td><td>άμύνω {V}</td><td>2</td><td>άνανεόομαι (-όω) {V}</td><td>-</td><td>άνίστημι {V}</td><td>3</td></t<>	ãyw {V}	53	άμύνω {V}	2	άνανεόομαι (-όω) {V}	-	άνίστημι {V}	3
(ξομαι (Ψ) 14 άμφιοβητέω (Ψ) 1 άναπαιώω (Ψ) 7 άνοιταίνω (Ψ) ω (Ψ) 1 άναβαίνω (Ψ) 3 άναπείθω (Ψ) 1 άνοιρώ (Ψ) ω (Ψ) 1 άναβαίνω (Ψ) 1 άναστάσου (Ψ) 1 άνασμώζω (Ψ) εία (Ψ) 1 άναγκάζω (Ψ) 4 άναπλήσου (Ψ) 5 άνατεγέρω (Ψ) μαι (Ψ) 1 άναγκάζω (Ψ) 1 άναστατή (Ψ) 4 άνατεισφέρω (Ψ) μαι (Ψ) 27 άναβείννημι (Ψ) 2 άναστατή (Ψ) 1 άνατεισφέρω (Ψ) ψη 5 άναδείννημι (Ψ) 1 άναστατή (Ψ) 1 άνατεισφέρω (Ψ) (Ψ) 5 άναδείση (Ψ) 1 άναστατή (Ψ) 1 άνατεισφέρω (Ψ) (Ψ) 5 άνασδόση (Ψ) 1 άνατείνημι (Ψ) 2 άνατερώ (Ψ) (Ψ) 22 άνατερώ (Ψ) 1 άνατρέμω (Ψ) 1 άνατρώμω (Ψ) (Ψ) 3 άνατρώμα (Ψ) 1 άνατρόμω (Ψ) 1 άνατρόμω (Ψ) (ψ) 3 άνατρόμα (Ψ) 1 άνατρόμα (Ψ) 1 άνατρόμα (Ψ) (ψ) 3 άνατρόμα (Ψ) 3 άνατρόμα (Ψ) 3 άνατρόμα (Ψ) (ψ) 1 άνασκολύττα (Ψ) 1 άναρφέρα (Ψ) 3 άνατρόμα (Ψ) (ψ) 2 άνασκολύττα (Ψ) 1 άναρφέρα (άγωνιάω {V}	2	άμφιέννυμι {V}	-	άνανεύω {V}	-	άνίσχω {ν}	-
ω (V) 3 φαπείθω (V) 3 φαπείθω (V) 1 φνοφάλω (V) ω (V) 34 φναβάλω (V) 1 φαπλάσουω (V) 1 φνοφήχω (V) εία (V) 11 φναγγγώσκω (V) 2 φναπληρώω (V) 4 φναπληρώω (V) εία (V) 11 φναγγγώσκω (V) 2 φναπλημώω (V) 4 φναπλημώω (V) ρία (V) 2 φναπλία (V) 1 φναπτώρη (V) 4 φναπτώρη (V) ρία (V) 2 φναπρέκονη (V) 1 φναπτώρη (V) 1 φναπτώρη (V) νω (V) 5 φναδρόρια (V) 1 φναπτώρη (V) 1 φναπρώτω (V) νω (V) 2 φναπρώτω (V) 1 φναπρώτω (V) 1 φναπρώτω (V) φ (V) 3 φναπρώτω (V) 1 φναπρώτω (V) 1 φναπρώτω (V) φ (V) 3 φναπρώτω (V) 1 φναπρώτω (V) 1 φναπρώτω (V) 2 φναπρώτω (V) φ (V) 3 φναπρώτω (V) φ (V) 1 φναπκολύτα (V) 2 φναπρώτω (V) 3 φναπρώτω (V) 3 φναπρώτω (V) 3 φναπρώτω (V) 3 φναπρώτω (V) φ (V) 2 φναπκολύτα (V) 3 φναπρώτω (V) 3 φναπρώτω (V) 3 φναπρώτω (V)	άγωνίζομαι {ν}	14	άμφισβητέω {V}	-	άναπαύω {V}	7	άνοηταίνω {V}	-
tω (V) 34 ἀναβάλλω (V) 1 ἀναγνάσκω (V) 1 ἀναγνάσκω (V) 1 ἀναγγάσκω (V) 1 ἀναγγάσκω (V) 1 ἀναγγάσκω (V) 2 ἀναπληρόω (V) 4 ἀναπληρόω (V) 4 ἀναπράρμω (V) 7 ἀναπράρμω (V) 7 ἀναπράρμω (V) 7 αναπράρμω (V) 9 ἀναθήδημω (V) 9 ἀναθήδημω (V) 1 αναπράρμω (V) 2 αναπράρμω (V) 2 αναπράρμω (V) 2 αναπράρμω (V) 3 αναπράρμω (V) <td>άγωνοθετέω {V}</td> <td>-</td> <td>άναβαίνω {V}</td> <td>3</td> <td>άναπείθω {V}</td> <td>-</td> <td>άνοίγω {V}</td> <td>-</td>	άγωνοθετέω {V}	-	άναβαίνω {V}	3	άναπείθω {V}	-	άνοίγω {V}	-
ξω (V) 1 άνστγγνώσκω (V) 4 άνστγληρώω (V) 5 άντετρέχομαι (V) ξω (V) 1 άνστβάω (V) 1 άνσττευφέρω (V) 4 άντετσέρχομαι (V) μαι (V) 27 άνάμω (V) 1 άνστετώω (V) 1 άντετισθέρχομαι (V) (V) 5 άναδέρχομαι (V) 1 άνστετώω (V) 1 άντετισθήμι (V) (V) 5 άναδέσχομαι (V) 1 άνστετώω (V) 1 άντετισθήμι (V) (V) 5 άναδέσχομαι (V) 1 άνστετίω (V) 1 άντετισθήμι (V) (V) 2 άναδέσχομαι (V) 3 άνατεβήμαι (V) 1 άντεξώμ (V) (V) 2 άνατεβήμαι (V) 1 άντεβλέπω (V) 1 άντεβλέμω (V) (V) 2 άνατρέμω (V) 1 άντεβλέμω (V) 1 άντεβλέμω (V) αν (V) 3 άνατρέμω (V) 1 άντεβλέμω (V) 1 άντεβλέμω (V) αν (V) 3 άναταθήμα (V) 1	άδικέω {V}	34	άναβάλλω (V)	-	άναπλάσσω {V}	-	άνοιμώζω {Λ}	-
ξω (V) 11 ἀναγκάζω (V) 2 ἀναπνέω (V) 4 ἀνετεσέρχομαι (V) ξω (V) 1 ἀναγράφω (V) 1 ἀνάπτω (V) 4 ἀνετεσέρχομαι (V) γ(V) 27 ἀνάδεἰκνυμι (V) 1 ἀναπείνω (V) 1 ἀναπείνω (V) (V) 5 ἀναδεἰκνυμι (V) 1 ἀναπείνω (V) 1 ἀναπείνω (V) (V) 5 ἀναδείσακω (V) 2 ἀναπείνω (V) 1 ἀναπείνω (V) (V) 22 ἀναδρίσακω (V) 3 ἀναπείνω (V) 1 ἀναπείνω (V) (V) 25 ἀναδρίπω (V) 1 ἀναπρέμω (V) 1 ἀναπρέμω (V) (Δ) 3 ἀναπρίσω (V) 1 ἀναπρέμω (V) 1 ἀναπρέμω (V) (Δ) 3 ἀναπράμω (V) 1 ἀναπρέμω (V) 1 ἀναπρέμω (V) (Δ) 1 ἀναπράμω (V) 2 ἀναπρίμω (V) 1 ἀναπρέμω (V) (Δ) 2 ἀναπράμω (V) 3 ἀναπρίμω (V) 3 ἀναπρίμω (V) (Δ) 1 ἀναπράμω (V) 2 ἀναπράμω (V) 3 ἀναπρίμω (V) (Δ) 2 ἀναπράμω (V) 3 ἀναπράμω (V) 3 ἀναπράμω (V) (Δ) 2 ἀναπράμω (V) 3 ἀναπράμω (V) 3 ἀναπράμω (V) (Δ) 2 ἀναπράμω (V) 3 ἀναπράμω (V) 3 ἀνα	άδοξέω { V }	-	άναγιγνώσκω {V}	4	άναπληρόω {V}	5	άντεγείρω {V}	-
ξω {V} 1 άναμτω {V} 4 άντεισφέρω {V} μαι {V} 27 άνάγω {V} 2 άνασξάγω {V} 1 άνατειών {V} 1 άνατειών {V} (V) 5 άναδξάγομαι {V} 1 άνατειών {V} 1 άνατειών {V} 1 άνατειών {V} (V) 22 άναδξάγωμαι {V} 2 άνατβήγαι {V} 2 άνατβήγαι {V} 1 άνατβάγω {V} (V) 25 άναδγητέω {V} 1 άνατβάγω {V} 1 άνατβάτω {V} 1 άνατβάτω {V} είν {V} 3 άνατβάν {V} 1 άνατβάτω {V} 1 άνατβάτω {V} είν {V} 3 άνατβάτω {V} 1 άνατβάτω {V} 1 άνατβάτω {V} είν {V} 3 άνατβάτω {V} 1 άνατβάτω {V} 1 άνατβάτω {V} είν {V} 3 άνατβάτω {V} 3 άνατβάτω {V} 3 άνατβάτω {V} είν {V} 1 άνακβάτω {V} 3 άνατβάτω {V} 3 άνατβάτω {V} είν {V} 2 άνακβάτα {V} 3 άνατβάτω {V} 3 άνατβάτω {V} είν {V} 2 άνακβάτα {V} 3 άνατβάτω {V} 3 άνατβάτω {V} είν {V} 2 άνακβάτα {V} 3 άνατβάτω {V} 3 άνατασθαίω {V} είν {V} 2 άνακβάτω {V} 3 άνατβάτω {V} 3 άνατασθαίω {V} <td>άειδω {V}</td> <td>11</td> <td>άναγκάζω {\\V}}</td> <td>2</td> <td>άναπνέω {V}</td> <td>4</td> <td>άντεισέρχομαι {V}</td> <td>-</td>	άειδω {V}	11	άναγκάζω {\\V}}	2	άναπνέω {V}	4	άντεισέρχομαι {V}	-
γ(γ) 27 ἀνάγω (γ) 2 ἀναστατόω (γ) 1 ἀντεπιδείκνυμι (γ) (V) 5 ἀναδείκνυμι (γ) 1 ἀναστατόω (γ) 1 ἀντεπιδείκνυμι (γ) (V) 5 ἀναδείκοικω (γ) 1 ἀναπείδημι (γ) 2 ἀντεμβείκνυμι (γ) (ν) 22 ἀναδίνομαι (γ) 3 άνατρέμω (γ) 1 ἀντερία (γ) γ (γ) 25 ἀναδίνεμα (γ) 1 ἀνατρέμω (γ) 1 ἀντερία (γ) γ (γ) 3 ἀνατρέμω (γ) 1 ἀνατρία (γ) 1 ἀντερία (γ) ν (γ) 32 ἀναποθαίτω (γ) 1 ἀνατρέμω (γ) 1 ἀντερία (γ) ν (γ) 32 ἀναποθαίτω (γ) 2 ἀνατράμο (γ) 1 ἀντερίω (γ) ν (γ) 1 ἀναπολύπτω (γ) 2 ἀναπρέμω (γ) 3 ἀντισταθμέω (γ) ν (γ) 1 ἀναπολύπτω (γ) 1 ἀναρότζω (γ) 3 ἀντισταθμέω (γ) ν (γ) 2 ἀναπολύπτω (γ) 3 άναρίζ	άθυμέω {V}	-	άναγράφω {V}	-	άνάπτω {V}	4	άντεισφέρω {V}	-
(γ) 5 ἀναδεξικυμι (γ) 1 ἀναστατόω (γ) 1 ἀναπτίθημι (γ) (γ) 5 ἀναδεξκομαι (γ) 1 ἀναπτίθημι (γ) 2 ἀντερῶ (γ) κοριαι (γ) 22 ἀναδιόριαι (γ) 3 άνατλῆγιαι (γ) 1 ἀντερῶ (γ) γ (γ) 25 ἀναδίοριαι (γ) 1 ἀνατρέπω (γ) 1 ἀντεβρίω (γ) ιών (γ) 1 ἀνατρέμω (γ) 1 ἀνατρίομω (γ) 1 ἀντιβάνω (γ) ω (γ) 3 ἀναποινίω (γ) 1 ἀνατρέμω (γ) 1 ἀντιβόμω (γ) ω (γ) 3 ἀναποινίω (γ) 2 ἀναπρέμω (γ) 3 ἀντιβόμω (γ) ω (γ) 1 ἀναποινίω (γ) 2 άναπρέμω (γ) 3 άντιδύω (γ) ω (γ) 1 ἀναποινίω (γ) 3 άντιδύω (γ) 3 άντισταθμέω (γ) ω (γ) 2 ἀναποινίω (γ) 3 άντισταθμέω (γ) 3 άντισταθμέω (γ) ω (γ) 2 άναποιδύπ (γ) 3 άντισταθμ	αίδέομαι {V}	27	ἀνάγω {ν}	2	άνασείω {V}	-	άντεξάγω {V}	-
{V} 5 ἀναδέχομαι {V} 1 ἀναπείνω {V} 1 ἀναπείνω {V} 1 ἀναπείνω {V} νουμαι {V} 22 ἀναδύσμαι {V} 3 ἀναπίθημι {V} 1 ἀντέχω {V} γ (V) 25 ἀναδύσμαι {V} 1 ἀναπρέπω {V} 1 ἀντήβαίνω {V} σριαι {V} 1 ἀναπρέμω {V} 1 ἀντιβλέπω {V} 1 ἀντιβλέπω {V} σριαι {V} 3 ἀναπρέω {V} 1 ἀντιβλέπω {V} 1 ἀντιβλέπω {V} π (V) 3 ἀναπρέω {V} 1 ἀντιβλέπω {V} 1 ἀντιβλέπω {V} π (V) 3 ἀναπρέω {V} 1 ἀντιβλέπω {V} 1 ἀντιβλέπω {V} π (V) 3 ἀναπρέω {V} 1 ἀντιβλέπω {V} 3 ἀντιδίδωμι {V} π (V) 3 ἀναπλύπτω {V} 3 ἀντισόω {V} 3 ἀντισόω {V} π (V) 2 ἀναφλύπτω {V} 3 ἀντισόμε {V} 3 ἀντισόμε {V} π (V) 2 ἀναφλύπτω {V} 3 ἀντισταθμέω {V} 3 ἀντισταθμέω {V}	aipéw {V}	2	άναδείκνυμι {V}	-	άναστατόω {V}	-	άντεπιδείκνυμι (V)	-
νομαι (V) 9 άναδιδάσκω (V) 2 άνατίθημι (V) 2 άνατίθημι (V) 1 άντέχω (V) νω (V) 25 άναδύσμαι (V) 1 άνατβαίνω (V) 1 άνατβαίνω (V) ομαι (V) 1 άναθάλτω (V) 1 άνατβάνω (V) 1 άντιβλέπω (V) κώω (V) 3 άναιρέω (V) 1 άνατβαμβάνω (V) 1 άντιλδήμα (V) ω (V) 32 άναικαθαίρω (V) 1 άνατβάνω (V) 1 άντιλέψω (V) α (V) 15 άναικαλλέω (V) 3 άναιξύω (V) 3 άντιστάμέω (V) νω (V) 2 άναικαλλέτω (V) 3 άντιστάμέω (V) 3 άντιστάμέω (V) νω (V) 2 άναικαλλέτω (V) 3 άντισταθμέω (V) 3 άντισταθμέω (V)	αἴρω {V}	2	άναδέχομαι {V}	-	άνατείνω {V}	-		-
νω {ν} 22 ἀναδύομαι {ν} 3 ἀνατλήναι {ν} 1 ἀνατρέτω {ν} ι {ν} 25 ἀναζητέω {ν} 1 ἀνατρέτω {ν} 1 ἀνατρέτω {ν} ομαι {ν} 9 ἀναθάλτω {ν} 1 ἀνατρέφω {ν} 1 ἀντιβλέτω {ν} νωθέω {ν} 3 ἀναιρέω {ν} 1 ἀναφέρω {ν} 2 ἀνατβίσμι {ν} νω {ν} 32 ἀνακαθαίρω {ν} 1 ἀναγρέω {ν} 1 ἀντιλάμβάνω {ν} νω {ν} 1 ἀνασκαινίζω {ν} 2 ἀναγωρέω {ν} 3 ἀντιδόμω {ν} νω {ν} 1 ἀνασκαινίζω {ν} 3 ἀναιφύχω {ν} 3 ἀντισόω {ν} νω {ν} 2 ἀνασκαλύπτω {ν} 1 ἀνδρίζω {ν} 3 ἀντισταθμέω {ν} νω {ν} 2 ἀνακαλύπτω {ν} 1 ἀνδρίζω {ν} 3 ἀντισταθμέω {ν}	αίσθάνομαι {V}	6	άναδιδάσκω {V}	2	άνατίθημι {V}	2	άντερῶ {V}	-
(V) 25 ἀναζητέω (V) 1 ἀνατρέτω (V) 1 ἀνατβαίνω (V) ομαι {V} 9 ἀναθάλτω {V} 1 ἀνατρέφω {V} 2 ἀντιβλέτω {V} ιὰω {V} 1 ἀνατρέω (V) 1 ἀνατβίωμι (V) 1 ἀντιβιών (V) ιω {V} 32 ἀναιρέω {V} 1 ἀνατλέμω {V} 1 ἀντιλέμω {V} ια {V} 1 ἀναταλέω {V} 2 ἀναιφύχω {V} 3 ἀντισόω {V} ια {V} 1 ἀνακαλύπτω {V} 1 ἀνδρίζω {V} 3 ἀντισταθμέω {V} ιω {V} 2 ἀνατισταθμέω {V} 3 ἀντισταθμέω {V}	αίσχύνω {V}	22	άναδύομαι {V}	8	άνατλήναι {ν}	-	άντέχω {V}	3
φμαι {V} 9 ἀναθάλττω {V} 1 ἀνατρέφω {V} 1 ἀνατβλέττω {V} τάω {V} 1 ἀναφέω {V} 2 ἀναφέω {V} 2 ἀντιδίδωμι {V} π ⟨V⟩ 32 ἀνακαθαίρω {V} 1 ἀναφέω {V} 1 ἀντιλέρω {V} π ⟨V⟩ 1 ἀνακανίζω {V} 2 ἀναγωρέω {V} 3 ἀντιδίω {V} π ⟨V⟩ 15 ἀνακαλέω {V} 3 ἀναφύχω {V} 2 ἀντισσθμέω {V} π ⟨V⟩ 2 ἀνακαλύπτω {V} 1 ἀνδρίζω {V} 3 ἀντισταθμέω {V} π {V} 2 ἀνακαλύπτω {V} 1 ἀνδρίζω {V} 3 ἀντισταθμέω {V}	aitéw {V}	25	άναζητέω {V}	-	άνατρέπω {V}	-	άντιβαίνω {V}	-
τάνατίνομαι {V} 1 ἀνατόρμαι {V} 2 ἀνατιδίδωμι {V} τοθέω {V} 3 ἀνατρέω {V} 1 ἀντιλαμβάνω {V} ω {V} 32 ἀναταθαίρω {V} 1 ἀντιλέγω {V} ποριαι {V} 1 ἀνατακαινίζω {V} 3 ἀντιτόρω {V} πο {V} 2 ἀνατασλέω {V} 3 ἀντισταθμέω {V} πο {V} 1 ἀνδρίζω {V} 3 ἀντισταθμέω {V} πο {V} 2 ἀνατασθμέω {V} 3 ἀντισταθμέω {V}	αίτιάομαι {V}	6	άναθάλπω {V}	-	άνατρέφω {V}	-	άντιβλέπω {V}	-
υυθέω (V) 3 ἀναιρέω (V) 1 ἀναιρβάνω (V) ω (V) 32 ἀνακαθαίρω (V) 1 ἀναιλέμω (V) ἀοριαι (V) 1 ἀναικαλέω (V) 3 ἀντισόω (V) υ (V) 1 ἀνδρίζω (V) 3 ἀντισταθμέω (V) ιω (V) 2 ἀναισάλύπτω (V) 1 ἀνδρίζω (V) 3 ἀντισταθμέω (V)	άκηδιάω {V}	-	άναίνομαι {\begin{align*} \begin{align*} ali	-	άναφέρω {V}	2	άντιδίδωμι {V}	80
ω {V} 32 ἀνακαθαίρω {V} 1 ἀνσιχέω {V} 1 ἀνσιλέγω {V} ἀοριαι {V} 1 ἀνακανίζω {V} 2 ἀνακαλέω {V} 3 ἀντισόω {V} υ {V} 15 ἀνακαλέπτω {V} 3 ἀντισταθμέω {V} 2 ἀντισταθμέω {V} κω {V} 2 ἀνακαλύπτω {V} 1 ἀνδρίζω {V} 3 ἀντισταθμέω {V}	άκολουθέω {V}	3	άναιρέω {V}	2	άναφύω {V}	-	άντιλαμβάνω {V}	2
άοραι {V} 1 ἀνακαινίζω {V} 2 ἀναχωρέω {V} 3 ἀντιξύω {V} υ {V} 15 ἀνακαλύπτω {V} 3 ἀναισταθμέω {V} 2 ἀντισταθμέω {V} ιω {V} 2 ἀνακαλύπτω {V} 1 ἀνδρίζω {V} 3 ἀντισταθμέω {V} Liste Temmes - Analyses 1 1 1	άκούω {V}	32	άνακαθαίρω {V}	-	άναχέω {V}	-	άντιλέγω {V}	9
υ {V} 15 ἀνακαλιέω {V} 3 ἀναιφύχω {V} 2 ἀντισόω {V} ιω {V} 2 ἀνακαλύπτω {V} 1 ἀνδρίζω {V} 3 ἀντισταθμέω {V} Liste Temmes - Analyses Liste Temmes - Analyses	άκροάομαι {V}	-	άνακαινίζω {ν}	2	άναχωρέω {V}	m	άντιξύω {V}	-
(V) 2 ἀνακαλύπτω (V) 1 ἀνδρίζω (V) 3 ἀντισταθμέω (V) Liste Temmes - Analyses' Liste Temmes - Analyses'	άλγέω {V}	15	άνακαλέω {V}	3	άναψύχω {align="center" (\begin{align=center" (\begin{align	2	άντισόω {V}	-
Liste 'Lemmes - Analyses'	άλείφω {V}	2	άνακαλύπτω {V}	-	άνδρίζω {V}	8	άντισταθμέω {V}	-
	21-11-17		Lis	te Lemme	is - Analyses'			Page 1 sur 13

Fig. 15. Alphabetial index of the verbs in the corpus of the letters by Gregory of Nazianzus

eipi {V}	462	εύρίσκω {V}	35	(Λ) πΣισπιλ	22	φθέγγομαι {V}	16
έχω { V }	286	άδικέω {V}	34	μηνήσκω (V)	22	φρονέω {V}	16
λέγω {V}	219	πράσσω {V}	34	παρίστημι {V}	22	άλγέω {V}	15
γίγνομαι {V}	180	έπιστέλλω {V}	33	Xpή {V}	22	άπολαύω {V}	15
ποιέω {V}	176	άκούω {V}	32	θεραπεύω {V}	21	εύφραίνω {V}	15
οίδα { V }	129	kpivw {V}	32	ύπολαμβάνω (V)	21	μέλλω {V}	15
δέω (δεήσω) {V}	128	θαυμάζω {\\V}	31	κινέω { V }	20	όμολογέω {V}	15
γράφω {V}	104	λυπέω {V}	31	κρατέω {V}	20	σιωπάω {V}	15
δίδωμι {V}	85	άξιόω {V}	30	ποθέω {V}	20	φεύγω {\\}	15
òpáw {V}	78	άγω {V}	29	φυλάσσω {\V}	20	άγωνίζομαι {Ν}	14
πάσχω {V}	73	άρχω {V}	29	άγνοέω {V}	19	<u>ذ</u> ش { ٧ }	14
βούλομαι {\\V\}	89	χράομαι (-άω) {V}	29	άπαιτέω {V}	19	κελεύω {V}	14
πάρειμι (είμί) {V}	29	ňкш {V}	28	έρχομαι {ν}	19	µévw ⟨V⟩	14
πείθω {V}	9	παρακαλέω {V}	28	ζάω { V }	19	παρέχω {V}	14
δοκέω {V}	62	πιστεύω {V}	28	κάμνω {V}	19	{ \(\) \(\) \(\) \(\) \(\) \(\) \(\) \(\)	14
φημί {V}	09	αίδέομαι (V)	27	λύω {V}	119	άπατάω {V}	13
τιμάω {V}	55	έθέλω {ν}	27	εύχομαι {ν}	18	άφίημι (V)	13
τυγχάνω {V}	25	έρῶ {V}	27	(Λ) ποιπίυ	18	διαλέγω {V}	13
φέρω {V}	52	διδάσκω {V}	26	μανθάνω {V}	18	καταγιγνώσκω {V}	13
δέχομαι (V)	51	aitéw {V}	25	βοηθέω (V)	17	καταλείπω {V}	13
έπαινέω {V}	20	έγκαλέω {V}	25	δηλόω {V}	17	κατέχω {V}	13
δύναμαι (V)	47	σπουδάζω {\\}	25	έλπίζω {Λ}	17	μέλω {V}	13
(Λ) σ)γίζο	47	Xaipu {V}	25	ήσσάομαι (-άω) {Ν}	17	μετέχω {V}	13
χαρίζομαι (-ω) {Λ}	47	βουλεύω {V}	24	ίστημι { V }	17	μιμέομαι {V}	13
γιγνώσκω {V}	46	€ouxa {V}	24	καταλύω {V}	17	όνομάζω {Λ}	13
painu {V}	46	προσάγω {V}	24	καταξιόω {V}	17	óppáw {V}	13
φιλοσοφέω {V}	46	καλέω {V}	23	τίθημι {V}	17	πληρόω {V}	13
λαμβάνω (V)	44	vukáw {V}	23	φροντίζω {ν}	17	προσαγορεύω {V}	13
θαρσέω {V}	43	συγχωρέω {V}	23	παίζω {Λ}	16	προστρέχω {V}	13
οῖομαι (-ω) {Λ}	45	αίσχύνω {ν}	22	πρεσβεύω {V}	16	συγγιγνώσκω {ν}	13
προστίθημι {V}	40	βλέπω {V}	22	πυνθάνομαι {V}	16	συνάγω {ν}	13

Fig. 16. Frequency index of the verbs in the corpus of the letters by Gregory of Nazianzus

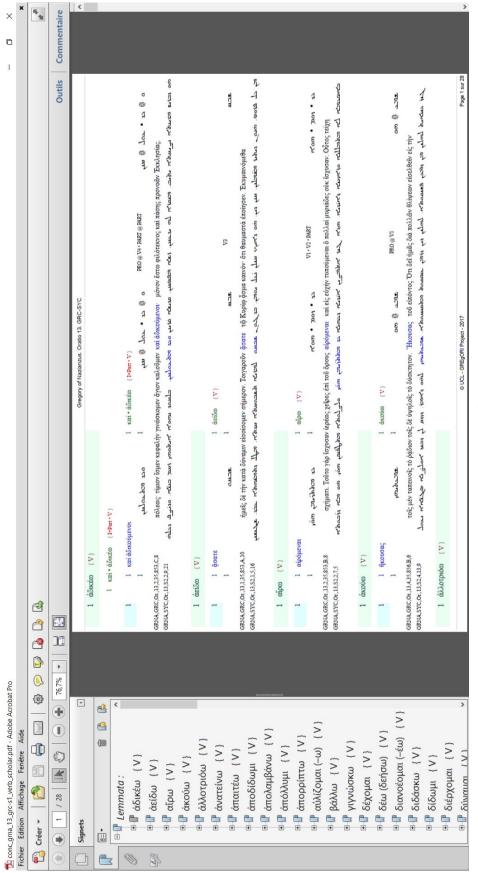


Fig. 17. Bilingual Greek-Syriac (S2) concordance of the 13th homily by Gregory of Nazianzus (scholar concordance)

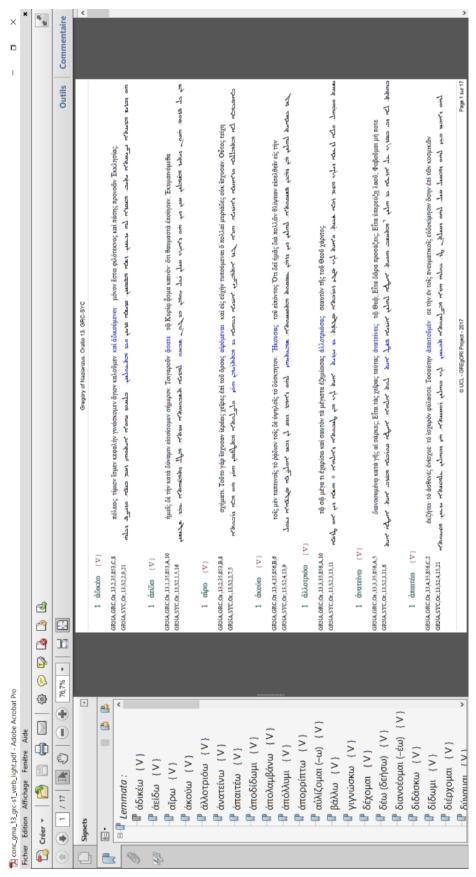


Fig. 18. Bilingual Greek-Syriac (S2) concordance of the 13th homily by Gregory of Nazianzus (light concordance)

CONCLUSION

For many years, all these technical developments and linguistic data are used to provide scholars with lemmatized concordances. These concordances were formerly published by Brepols Publishers on microfiches in the *Thesaurus Patrum Graecorum* series (see [Coulie, 1996; Kindt, 2010]. This use of microfiches is however rightly characterized as a "doubtful aspect" since [Trapp, 2008:97]. Now concordances and others lexical outputs are directly available on the website of the GREgORI project.

Some collaborators made use of these computerized resources to build lemmatized indexes (*index nominum*, *index verborum*, etc.) included in their editions of ancient texts (see for instance [Schmidt, 2001], [Auwers, 2005], [Sanpeur, 2007], [Amato, 2009], [Auwers, 2011] and more recently [Calzolari, 2017]). The GREgORI project also takes part in international research programs, such as the *Syriac Galen Palimpsest* project led by the University of Manchester or the study undertaken at the University of Lausanne and devoted to the comparative analysis of the *Iliad* and its Byzantine translation contained in the manuscript Genavensis 44.

To date, a web-based interface allowing to consult lemmatized concordances on personal computers via Internet, is being prepared in collaboration with the Cental. A beta-version of this interface is used by Peeters Publishers (Leuven) in order to pave the path for a forthcoming digital version of the well-known *Corpus Scriptorum Christianorum Orientalium*, the series dealing with some of the most important ancient sources written in languages of the Christian Middle East.

References

Amato E. Severus Sophista Alexandrinus Progymnasmata quae exstant omnia [...] accedunt Callinici Petraei et Adriani Tyrii sophistarum testimonia et fragmenta necnon incerti auctoris ethopoeia nondum vulgata. Bibliotheca Scriptorum Graecorum Romanorum Teubneriana, Walter de Gruyter (Berlin, New York), 2009.

Auwers J.-M. Concordance du Siracide (Grec II et Sacra Parallela), avec la collaboration d'Églantine Proksch-Strajtmann. Cahiers de la Revue Biblique, 58. Gabalda (Paris), 2005.

Auwers J.-M. Procopii Gazaei Epitome in Canticum Canticorum, cum praefatione a Jean-Marie Auwers et Marie-Gabrielle Guérard curata. Corpus Christianorum. Series Graeca, 67. Brepols Publishers (Turnhout), 2011.

Calzolari, V. Apocrypha Armeniaca. Acta Pauli et Theclae, Prodigia Theclae, Martyrium Pauli. Corpus Christianorum. Series Apocryphorum, 20. Brepols Publishers (Turnhout), 2017.

Coulie B. La lemmatisation des textes grecs et byzantins: une approche particulière de la langue et des auteurs. *Byzantion*. 1996:66:35-54.

Coulie B., Kindt B. and Pataridze T. Lemmatisation automatique des sources en géorgien ancien. *Le Muséon*. 2013;126:161-201

Fleury S. mkAlign, Manuel d'utilisation, Université Sorbonne Nouvelle Paris 3, juin 2012. [get PDF]

Gallay P. and Jourjon M. Lettres théologiques. Sources Chrétiennes, 280. Les Éditions du Cerf (Paris),1976.

Gallay P. Saint Grégoire de Nazianze, Lettres. Collection des Universités de France. Les Belles Lettres (Paris), 1964-1967.

Kindt B. Des concordances et du *Thesaurus Patrum Graecorum*: la ligne éditoriale du Projet de recherche en lexicologie grecque (1990-2010). Schmidt A. *Studia Nazianzenica* II. Corpus Christianorum. Series Graeca. 73. Corpus Nazianzenum, 24. Brepols Publishers (Turnhout), 2010:43-120.

Kindt B. La lemmatisation des sources patristiques et byzantines au service d'une description lexicale du grec ancien. Les principes de formulation des lemmes du Dictionnaire Automatique Grec (D.A.G.). *Byzantion*. 2004;74:213-272. [get PDF]

Liddell H.G., Scott R. and Jones H.S. A Greek-English Lexicon, 9th ed, Oxford University Press. Clarendon Press (Oxford), 1940, repr. 1977 with Barber E. Greek-English Lexicon. A Supplement, Oxford University Press. Clarendon Press (Oxford), 1968; see also Glare G. W. Greek-English Lexicon. Revised Supplement, Oxford University Press. Clarendon Press (Oxford), 1996.

Pataridze T. and Kindt B. Text Alignment in Ancient Greek and Georgian: A Case-Study on the First Homily of Gregory of Nazianzus. *Journal of Data Mining and Digital Humanities*. 2017;Special Issue.

Paumier S. Unitex 3.1 User Manual, Université Paris-Est Marne-la-Vallée, March 27, 2016. [get PDF]

Sanspeur Cl. Sancti Gregorii Nazianzeni Opera. Versio armeniaca, IV. Oratio VI. Corpus Christianorum. Series Graeca, 61. Corpus Nazianzenum, 21. Brepols Publishers (Turnhout), 2007.

- Schmidt A.B. Sancti Gregorii Nazianzeni Opera. Versio Syriaca. II. Orationes XIII, XLI. Corpus Christianorum. Series Graeca, 47; Corpus Naizanzenum, 15. Brepols Publishers (Turnhout, Leuven), 2002.
- Schmidt Th. Basilii Minimi in Gregorii Nazianzeni Orationem xxxviii Commentarii. Corpus Christianorum. Series Graeca, 46. Corpus Nazianzenum. 13. Brepols Publishers (Turnhout, Leuven), 2001.
- Trapp E. Lexicography and electronic textual resources. Jeffreys E., Haldon J.F., Cormack R. (eds) The Oxford Handbook of Byzantine Studies. Oxford University Press (Oxford), 2008:95-100.
- Tuerlinckx L. La lemmatisation de l'arabe non classique. Dister A., Fairon C. and Purnelle G. (eds) *Le poids des mots. 7^{es} Journées internationales d'Analyse statistique des Données Textuelles, 10-12 mars 2004, Louvain-la-Neuve*. Vol 2. Cahiers du Cental, II. Presses Universitaires de Louvain (Louvain-la-Neuve), 2004:1069-1078. [get PDF]
- Van Elverdinghe E. Recurrent Pattern Modelling in a Corpus of Armenian Manuscript Colophons. *Journal of Data Mining and Digital Humanities*. 2017;Special Issue.