

25 Electronic excerption of lemmas in Modern Greek

John N. Kazazis, Rania Voskaki

ABSTRACT

The current research regards the electronic excerption of lemmas in Modern Greek in synchronic terms. The ultimate goal is to create a basis for further linguistic study through various potential applications. The project lies within the framework of D.J. Georgacas' archive enhancement. Given that his archival material is gradually digitised and registered in a database, we opted for continuing with its enhancement by creating slips in electronic form ab initio. Electronic slips are saved in two forms: in a machine-processable form (.doc) and in an image representation form (.jpg). Currently, the created material amounts to 25,000 electronic slips.

As for the methodology and the excerption criteria, we follow the guidelines set forth by D.J. Georgacas. The electronic slips contain a brief excerpt of written discourse involving the lemma that has been selected to be indexed. The extent of the excerpt depends on each lemma, with the aim of revealing its specific usage. In parallel, the electronic slip records the source, the publication date and the lemmatic form. We systematically excerpt citations from Modern Greek sources, originated in literature, arts, science and conversational language. The selected material has been published from 2000 to date.

Ηλεκτρονική αποδελτίωση λημμάτων της Νέας Ελληνικής

Ιωάννης Ν. Καζάζης, Ράνια Βοσκάκη

ΠΕΡΙΛΗΨΗ

Αντικείμενο της παρούσας έρευνας είναι η ηλεκτρονική αποδελτίωση λημμάτων της Νέας Ελληνικής σε συγχρονικό επίπεδο, με στόχο να αποτελέσει τη βάση για περαιτέρω μελέτη της γλώσσας μέσω ποικίλων πιθανών εφαρμογών. Εντάσσεται στο πλαίσιο εμπλουτισμού του αρχαιακού υλικού του Δ.Ι. Γεωργακά. Δεδομένου ότι το αρχαιακό αυτό υλικό, το οποίο αριθμεί 2.500.000 δελτία, ψηφιοποιείται σταδιακά και καταγράφεται σε βάση δεδομένων, αποφασίσαμε να συνεχίσουμε τον εμπλουτισμό δημιουργώντας εξ αρχής τα δελτία σε ηλεκτρονική μορφή. Αποθηκεύονται σε δύο μορφές: σε μορφή επεξεργάσιμου αρχείου (.doc) και σε μορφή εικόνας (.jpg). Αυτή τη στιγμή τα ηλεκτρονικά δελτία ανέρχονται στις 25.000.

Ως προς τη μεθοδολογία και τα κριτήρια αποδελτίωσης, ακολουθούμε τα πρότυπα που τέθηκαν από τον Δ.Ι. Γεωργακά. Αποδελτιώνουμε συστηματικά τις πηγές της Νέας Ελληνικής, από τη λογοτεχνία, τις τέχνες, την επιστήμη και την καθημερινή γλώσσα. Πιο συγκεκριμένα, τα παραθέματα αντλούνται από ένα ευρύ φάσμα γραπτών πηγών: α. εφημερίδες και περιοδικά, β. ανθολογίες ποίησης και πεζογραφίας, γ. χωριστές εκδόσεις συγγραφέων, ποιητών και πεζογράφων, δ. συλλογικούς τόμους, ε. επιστημονικά

άρθρα, στ. ειδικές μελέτες για την παιδεία, τις πολιτικές επιστήμες, τη φιλοσοφία και τις καλές τέχνες και ζ. γλωσσάρια. Το υλικό που επιλέγεται έχει δημοσιευθεί από το 2000 ως σήμερα.

Ενδεικτικά, αποδελτιώνουμε: 1. Ορολογία – τεχνικούς όρους, π.χ.: *πελίκη*, 2. Χλωρίδα – πανίδα, π.χ.: *γεοκάριο*, 3. Συμφράσεις, π.χ.: *καθεστηκυία τάξη*, 4. Εκφράσεις λόγιες και δημοτικές, π.χ.: *διεξάγω πόλεμο, το παίζω ζόρικος/η*, 5. Απολιθώματα, π.χ.: *εν γνώσει*, 6. Δάνειες λέξεις, π.χ.: *τυρί κόρατζ*, 7. Σπάνιες λέξεις, π.χ.: *μακρόφεγγος, -η, -ο*, 8. Υποκοριστικά – μεγεθυντικά, π.χ.: *χαμογελάκι, ταλεντάρα*, 9. Ουσιαστικοποιημένα επίθετα, π.χ.: *ξυλογλυπτική*, 10. Νεολογισμούς, π.χ.: *τζανκοτροφή*, 11. Συνεκδοχές, π.χ.: *χιλιάδες κεφάλια (πρόβατα)*, 12. Μεταφορές, π.χ.: *διαξιφισμοί*, 13. Μετωνυμίες, π.χ.: *το υπουργείο ανακοίνωσε*, 14. Κοινές λέξεις ως τοπωνύμια, π.χ.: *Πετροκέρασα*. Επίσης, αποδελτιώνουμε: προθηματοποιημένα ρήματα, μετοχές, ουσιαστικά με λημματικό τύπο στον πληθυντικό αριθμό, υποστηρικτικά ρήματα, κατηγορηματικά ονόματα, φράσεις, παροιμίες και λήμματα που προέρχονται από διαλέκτους.

Τα ηλεκτρονικά δελτία περιλαμβάνουν ένα σύντομο παράθεμα γραπτού λόγου που εμπεριέχει το λήμμα που κρίνεται ενδιαφέρον να αποδελτιωθεί. Η έκταση του παραθέματος ορίζεται κατά περίπτωση με σκοπό να αποτυπώσει τη χρήση του υπό εξέταση λήμματος. Παράλληλα, στο ηλεκτρονικό δελτίο αναγράφεται η πηγή, η χρονολογία δημοσίευσης και ο λημματικός τύπος. Ο εκάστοτε λημματικός τύπος συνοδεύεται από μορφολογικό σχολιασμό, με τη μορφή αναρτημάτων σχολιασμού. Αξίζει να σημειωθεί το γεγονός ότι η γραμματική κατηγορία ορίζεται βάσει της χρήσης του λήμματος εντός του συγκεκριμένου συγκειμένου στο οποίο εντοπίζεται.

Ως προς τον μορφολογικό σχολιασμό, τα αναρτήματα σχολιασμού είναι ενδεικτικά τα εξής: 1. *indecl* για τα άκλιτα ουσιαστικά και επίθετα, π.χ.: *το σουξέ, έξτρα*, 2. *verbal adj* για τα ρηματικά επίθετα, π.χ.: *ανυφαντός, -ή, -ό*, 3. *adv* για τα επιρρήματα, π.χ.: *επιτόπου*, 4. *adv phr* για τις επιρρηματικές φράσεις, π.χ.: *εξ ευωνύμων*, 5. *adv use* για τις επιρρηματικές χρήσεις, π.χ.: *το πρωί*, 6. *pron* για τις αντωνυμίες, π.χ.: *τούτος, -η, -ο*, 7. *prep* για τις προθέσεις, π.χ.: *από*, 8. *excl* για τα επιφωνήματα, π.χ.: *σύρμα*, 9. *pref* για τα προθήματα, π.χ.: *συν-*, 10. *suff* για τα επιθήματα, π.χ.: *-άρος*, 11. *1st me cpds* για το πρώτο συνθετικό λέξεων, π.χ.: *μαστρο-*, 12. *part.* για τα μόρια, π.χ.: *γιατί*, 13. *conj* για τους συνδέσμους, π.χ.: *και*, 14. *abbr* για τις συντομογραφίες, π.χ.: *ΚΕΓ*, 15. *arithm* για τα αριθμητικά, π.χ.: *ένας, μία, ένα*, 16. *o/p* για τις ονοματοποιίες π.χ.: *πικ τακ*.

Το υλικό των ηλεκτρονικών δελτίων δεν συμπεριλαμβάνει κοινές λέξεις (π.χ. *ουρανός*), εκτός από τις περιπτώσεις εκείνες, όπου αποτελούν μεταφορά, μετωνυμία, συνεκδοχή ή μέρος φράσης. Κατά συνέπεια, το υλικό αποτελείται από λιγότερο συνηθισμένες λέξεις. Μελετώντας το σύνολο των ηλεκτρονικών δελτίων διακρίναμε τις ακόλουθες ενδεικτικές κατηγορίες: 1. Σπάνια λήμματα – σπάνιες χρήσεις λημμάτων, π.χ.: *ένα πανσέληνο, ασημένιο δολάριο, άξερτος*, 2. Συμφράσεις ορολογίας, π.χ.: *κληματίς η ελισαβέτιος-καρόλιος*, 3. Παραλλαγές στο επίπεδο του λόγου, π.χ.: *ηλεκτρονικός υπολογιστής-κομπιούτερ-κομπιούτορας*, 4. Μορφολογικές παραλλαγές, π.χ.: *ρεπροντυξιδόν-ρεπροντυξιδόν*, 5. Ιδιωματικές φράσεις, π.χ.: *κατεβάζω/τρώω τον αγλέουρα*, 6. Υποστηρικτικές δομές, π.χ.: *κάνω αίτηση*.

Τα ηλεκτρονικά δελτία που δημιουργήθηκαν αποτελούν ένα σημαντικό γλωσσικό πόρο, που απεικονίζει τη συγχρονία της Νέας Ελληνικής και σε συνδυασμό με το αρχαιακό υλικό του Δ.Ι. Γεωργακά επιτρέπει τη διαχρονική μελέτη της ελληνικής γλώσσας. Η συστηματική αποδελτίωση σύγχρονου γραπτού λόγου μπορεί να αποτελέσει τη βάση για περαιτέρω έρευνα, όπως τη μελέτη της εξέλιξης της Νέας Ελληνικής, τη χρήση του αποδελτιωμένου υλικού στο πλαίσιο συγγραφής ενός λεξικού (έντυπου ή/και ηλεκτρονικού) βασισμένο σε αυθεντικές πηγές, τη διδακτική αξιοποίηση για τη διδασκαλία της Νέας Ελληνικής ως δεύτερης/ξένης γλώσσας. Βασικός μας στόχος είναι ο συνεχής εμπλουτισμός των ηλεκτρονικών δελτίων.

0 Introduction

D.J. Georgacas' archive comprises 2.5 million slips, illustrating Modern Greek words. Chronologically, it contains Modern Greek words from the 18th to the 20th century, with greater bulk of words coming from the second half of the 20th century. It also preserves important words and specialised terms from the past, dating back to the 15th century or even earlier [2]. In 1996 it was bequeathed to the Centre for the Greek Language (CGL) which undertook its preservation and scholarly development.

The digitisation of the entire archival material, all in hardcopy, is being developed the past few years under the aegis of the CGL. Currently, the bulk of the slips has already been scanned and thereby preserved in electronic form. In addition, the digitised material is being inserted into a database, especially created to allow the morphological annotation of the lexical entries.

It is worth mentioning that from 1996 to 2007, the CGL continued to enhance the archive with a supplement of conventional slips. That material amounts 23,000 slips. However, it was proved to be laborious and time-consuming, given that it also had to be digitised. It is under this scope that the enhancing of the D.J. Georgacas archive by means of digital excerption ab initio was considered to be an imperative task.

1 Methodology

1.1 Corpus selection

The material employed for citations is gleaned from a representative spectrum of various textual sources: (a) newspapers and periodicals, (b) prose and poetry anthologies, (c) editions of single authors of poetry and prose, (d) collective volumes, (e) scientific papers, (f) scholarly works on education, political science, philosophy, and the fine arts, and (g) glossaries. The entire citation material dates from 2000 to date. The main objective when selecting the excerption corpus is the representativeness of the corpus in terms of textual genres.

There are no ad hoc examples of usage provided; each example has been drawn from the full range of attested Modern Greek usage. That is, every lemma is documented by phrases and passages excerpted from texts written by Modern Greek writers. It is therefore a representative corpus comprising the language of simple prose, artistic literature and poetry, historical works, essays, and the language of scholarly and scientific demotic Greek.

1.2 Excerption criteria

Regarding the excerption criteria we have adopted the standards and specifications established by D.J. Georgacas [5]. Morphological, semantic, and syntactic variations are systematically excerpted. An indicative list of the categories we consider during excerpting is the following:

1. Terminology – technical terms [8] [15]: *πελίκη* (pelike).
2. Flora and fauna: *γεοκάρυο* (geocaryum).
3. Collocations [4]: *καθεστηκυία τάξη* (status quo).
4. Literary and popular expressions: *διεξάγω πόλεμο* (to wage war),
το παίζω ζόρικος/η (to be a hard nut to crack).
5. Fossil words: *εν γνώσει* (knowingly).
6. Loan words: *τυρί κότατζ* (cottage cheese).
7. Rare words: *μακρόφρεγγος, -η, -ο* (gleaming far).
8. Diminutive and augmentative words: *χαμογελάκι* (soft smile),
ταλεντάρια (super talented).
9. Nominal adjectives: *ξύλογλυπτική* (wood carving).
10. Neologisms [1]: *τζανκοτροφή* (junk food).
11. Synecdoches [13]: *χιλιάδες κεφάλια (πρόβατα)* (thousands of heads (of sheep)).
12. Metaphors [10]: *διαξιφισμοί* (altercation).

13. Metonymies [9]: το υπουργείο ανακοίνωσε (the ministry announced).
14. Common words used as toponyms: Πετροκέρασα (sweetheart cherry & Petrokerasa: village in Greece).

In addition we excerpt: prefixed verbs, participles, nouns only in plural form, support verbs [6] [7], predicative nouns [6], phrases, proverbs¹ and lemmas originated from dialects [3]. Nevertheless, common words² are rarely excerpted, except for those which are part of a phrase³ or constitute a metaphor, a metonymy or a synecdoche. Considering that the bulk of the archival material comprises a set of slips for every lemmatic form and that common words had been thoroughly excerpted, we opted for focusing on the above mentioned categories.

1.3 Electronic form

Currently, the corpus created amounts to 25,000 electronic slips. Each slip contains a brief excerpt of written discourse involving the lemma that has been selected to be indexed. The extent of the excerpt depends on each lemma, with the aim of elucidating its specific usage. More precisely, the excerpt can range from a sentence to a paragraph.

As regards the form of each excerpt slip, data are presented as follows:

- (a) Centered on the slip appears the citation including the lemma chosen to be indexed; it is highlighted (underlined and in bold characters) in order to be easily detected while studying. That is, recording individual words and collocations within their respective context of occurrence onto excerpts slips.
- (b) On the left header we indicate the citation source⁴ as well as the publication date.
- (c) On the right header, we indicate the lemmatic form, followed by its grammatical category and a morphological annotation tag. That means that declined or conjugated words are recorded in nominative singular or in infinitive.

Moreover, the definition of the grammatical category depends on the specific usage of the lemma within the bilateral context. For example, the word *δωρυφόρος* (satellite) is a noun. Nevertheless, in the following excerpt it is an adjective:

¹ As for phrases and proverbs, we create as many slips as their lexical components.

² By common words, we refer for example to words like the following: *ουρανός* (sky), *αυτοκίνητο* (car), etc.

³ For example: *καθαρός ουρανός αστραπές δεν φοβάται* (a clear conscience fears no accusations).

⁴ In literary works we also add the page of the book where the lemma was detected.

epta imeres kathimerini 2003	δορυφόρος adj
<p>Κεντρικός οικισμός στην Υστερη Εποχή του Χαλκού και στην Πρώιμη Εποχή του Σιδήρου αναδεικνύεται ο οικισμός στην οχυρή θέση «Καστρί», την πληθυσμιακή άνθηση του οποίου μαρτυρούν οι εκτεταμένες νεκροπόλεις του και μικροί δορυφόροι οικισμοί γύρω του. Οι συστάδες των κτιστών οικογενειακών τάφων αντικατοπτρίζουν την κατά γέννη οργάνωση μιας κοινωνίας γεωργών και κτηνοτρόφων, η οποία έχει αναπτύξει δομές εξουσίας που της παρέχουν ασφάλεια σε τειχισμένες ακροπόλεις και πύργους.</p>	

Figure 1: Electronic slip example

The morphological annotation tags [2] used for each entry are as follows:

- (1) Verbs are mainly not annotated, except defective verbs that include an annotation tag illustrating the mood and the person; for example, *άσε* (let), *2sgimper*⁵.
- (2) Nouns are followed by the appropriate article. Moreover, for the indeclinable nouns we add the tag *indecl*; for instance, *σουξέ* (success), *το* (the), *indecl* (for indeclinable)⁶.
- (3a) For adjectives, we indicate the respective suffix for the feminine and the neuter form. In case of a foreign word, that has not been adapted to the Modern Greek morphological system, we use the annotation mark *adj* (for adjective) and *indecl*; for example, *έξτρα* (extra), *adj, decl*. (3b) Verbal adjectives are annotated by the tag *verbal adj*; for instance *ανυφαντός, -ή, -ό* (woven), *verbal adj*.
- (4a) As for adverbs, the tag *adv* is used as an annotation mark; for example, *επιόπου* (on the spot/immediately), *adv*. (4b) In case of an adverbial phrase, we add the tag *adv phr*; for instance, *εξ ευωνύμων* (on the left hand side). (4c) In case of an adverbial

⁵ *2sgimper* stands for second person, imperative mood.

⁶ Nominal adjectives are annotated as nouns.

- use, we insert the tag *adv use*; for example, *πρωί* (morning), *το* (the), *adv use*.
- (5) Pronouns are annotated with the tag *pron*; for instance, *τούτος*, *-η*, *-ο* (this), *pron*.
 - (6) Prepositions are annotated with the tag *prep*; for example, *από* (from), *prep*.
 - (7) Exclamations are followed by the tag *excl*; for instance, *σύρμα* (look out), *excl*.
 - (8) For prefixes we add the tag *pref*; for example, *συν-* (com-), *pref*.
 - (9) As for suffixes we use the tag *suff*; for instance, *-άρος*⁷, *suff*.
 - (10) The first member of compounds are annotated with the tag *1st me cpds*; for example, *μαστρο-*⁸, *1st me cpds*.
 - (11) Particles are annotated with the tag *part*.⁹; for instance, *γιατί* (why), *part*.
 - (12) As regards conjunctions, we add the tag *conj*; for example, *και* (and), *conj*.
 - (13) Abbreviations are annotated by the tag *abbr*; for instance *ΚΕΓ* (CGL), *abbr*.
 - (14) As for numbers, we use the tag *arithm*; for example, *ένας*, *μία*, *ένα* (one), *arithm*.
 - (15) Onomatopoeias are annotated by the tag *o/p*; for instance, *τικ τακ* (tick), *o/p*.

We have opted for saving the electronic slips in two forms: in a machine-processable form (.doc) and in an image representation form (.jpg). The first form allows the flexibility for future modifications in the context of potential applications, while the second form is in accordance with the digitised version of the archival material. Considering that the new excerption material will shortly be added to the archive, a homogeneous electronic form had to be provided.

2 Observation outline

As pointed out above (in 1.2), excerption of common words was not deemed necessary during the enhancement of the archival material. As a result we came up with a corpus consisting of uncommon words that we subsequently studied. Thereby, an overall study of

⁷ It is an augmentative noun suffix in Modern Greek. There is no equivalent suffix in English.

⁸ It constitutes the first member of a compound noun used in order to describe someone's ability as a technician.

⁹ We opted for adding a dot at the end of the tag, so as to avoid the ambiguity existing with the English word *part*.

the electronic corpus of 25,000 slips created, allowed the deduction of the following sub-categories:

1. Rare lemmas – rare usage of lemmas

Rare lemmas and rare usage of lemmas constitute a significant part of the electronic corpus of slips. As rare we define lemmas or usage of lemmas that are not registered in any Modern Greek dictionary. Additionally, most of them do not return any web search result. The textual sources of these lemmas are mostly citations in literary and journalistic texts. For example:

- a. ένα πανσέληνο, ασημένιο δολάριο (a perfectly round silver coin),
- b. αεροπνίχτης (dangerous aeroplane),
- c. άξερτος, -η, -ο (unknown).

2. Terminology collocations

Likewise, an important part of the corpus comprises terminology collocations. They are excerpted from thematic inserts in newspapers, from glossaries, and collective volumes. The number of electronic slips, created for each terminology collocation, depends on the number of words that each term is composed of. That is, a terminology collocation consisting of an adjective and a noun is excerpted in two separate slips; the first one refers to the lemmatic form of the adjective, while the second one refers to the respective lemmatic form of the noun¹⁰. For example:

- a. λιαστό κρασί ή οίνος πάσσος (raisin wine),
- b. κιρρός/ωχρός/μελιχρός οίνος (honey-coloured wines),
- c. κενταύρια η θρακική (Centaurea thracica¹¹),
- d. κληματίς η ελισαβέτιος-καρόλιος (Clematis elisabethae-carolae).

3. Register variations

The electronic corpus records variations across registers. Obviously, register variations are excerpted from all the recorded textual sources and genres. For example:

ηλεκτρονικός υπολογιστής–κομπιούτερ–κομπιούτορας (computer).

4. Morphological variations

In addition, morphological variations are systematically noted down. For example:

¹⁰ Articles that are part of the terminology collocation are not separately excerpted; as it is the case in the examples: c and d).

¹¹ We opted for a Latin rendering of the specific terms.

ρεπροντυξιόν–ρεπροντιξιόν (reproduction),

ταχιά–ταχιό (tomorrow).

5. Idiomatic phrases

Idiomatic phrases are subclass of non-free (or set/fixed/frozen) expressions. They constitute a multiword expression whose meaning cannot be deduced neither by the general rules of the language in question, nor by the meaning of the constituent words¹² [12]. For example:

κατεβάζω/τρώω τον αγλέουρα (to eat excessively).

6. Support verb constructions¹³

The notion of support verb construction denotes combinations of predicative nouns or adjectives and semantically weak or reduced verbs. Within that type of constructions the noun or adjective subcategorises semantically, and the verb syntactically [11]. For example:

κάνω αίτηση (submit an application).

The most numerous verb constructions—in terms of slips—throughout the electronic corpus contain the following support verbs: *κάνω* (to do), *έχω* (to have), *είμαι* (to be), and *γίνομαι* (to become). It is worth mentioning that in the case of support verb constructions, we create two slips: the first one for the verb and the second one for the noun or adjective; it is according to the guidelines we follow in the framework of electronic excerption.

3 Conclusions and prospects

The electronic slips that we have created the past six years constitute considerable material contributing to the representation of Modern Greek in synchronic terms. Moreover, along with the D.J. Georgacas archive it represents a significant corpus in diachronic terms. Therefore, the integration of the electronic slips in the database containing the archival material is among our priorities. The systematic excerption of written discourse can serve as a base for further research on the evolution of Modern Greek.

The integration of the electronic corpus of slips will allow thorough annotation; that constitutes the basis for accurate word recognition, essential for numerous applications. As for instance, further processing through semantic interconnection among lemmatic forms could be established in the framework of semantic networks.

¹² Idiomatic phrases are also determined as *phrasemes* [12].

¹³ Cf. [6] [7] [11] [14].

The systematically annotated corpus could serve as a base for lexicographic research, as it offers authentic examples of word usage. That is, for the construction of Modern Greek dictionaries in printed form, as well as for the enhancement of electronic dictionaries; the variety of different word forms that has been registered could be proven to be useful in computational lexicography.

In parallel, it could be used in teaching/learning Modern Greek, as it comprises a corpus of authentic written discourse. The alignment of literary citations could also be possible, considering that the literary works chosen to be excerpted are mostly translated into several languages.

Nevertheless, the existing electronic corpus of slips has to be constantly enhanced; that constitutes our main objective.

4 References

- [1] Αναστασιάδη-Συμεωνίδη, Ά, *Η νεολογία στην κοινή νεοελληνική*, Θεσσαλονίκη: Διδακτορική διατριβή, Α.Π.Θ, 1986.
- [2] Centre for the Greek Language, *A modern Greek-English dictionary*, compiled and edited by Demetrius J. Georcacas, published by Aristide D. Caratzas, New York & Athens, 2005, Volume 1.
- [3] Centre for the Greek Language, *La langue grecque et ses dialectes*, Athens, 2000.
- [4] Firth, J.R., “Modes of Meaning”, *Papers in Linguistics 1934-51*, London: Oxford University Press, pp. 190-215.
- [5] Georgacas, D.J., “The first large international dictionary of common and cultivated Modern Greek”, *Μαντατοφόρος 18*, 1981, pp. 35-66.
- [6] Giry-Schneider, J., *Les prédicats nominaux en français. Les phases simples à verbe support*, Genève, Droz, 1987.
- [7] Gross, M., « Les bases empiriques de la notion de prédicat sémantique », *Langages 63*, Paris : Larousse, 1981, pp. 7-52.
- [8] ISO 1087-1:2000, *Terminology work – Vocabulary – Part 1: Theory and application*.
- [9] Kövecses, Z., *Metaphor, a practical introduction*, New York: Oxford University Press, 2002.
- [10] Lakoff, G, “The contemporary theory of metaphor”, *Metaphor and Thought*, edited by Andrew Ortony, Cambridge: Cambridge University Press, 1993, pp. 202-251.
- [11] Langer, S., “A Formal Specification of Support Verb Constructions”, S. Langer, & D. Schnorbusch, *Semantik im Lexikon*, Tübingen: Narr, pp. 179-202.
- [12] Mel’čuk, I., “Phrasemes in language and phraseology in linguistics”, M. Everaert, E.-J. van der Linden, A. Schenk and R. Schreuder (eds.), *Idioms: Structural and psychological perspectives*, Hillsdale, NJ: Lawrence Erlbaum, 1995, pp. 167–232
- [13] Seto, K., “Distinguishing metonymy from synecdoche”, *Metonymy in Language and*

Thought, edited by Klaus-Uwe Panther & Günter Radden, Amsterdam/Philadelphia: John Benjamins, 1999, pp. 91-120.

- [14] Κλαίρης Χ. & Γ. Μπαμπινιώτης *Γραμματική της Νέας Ελληνικής, Δομολειτουργική-Επικοινωνιακή*, Αθήνα: Ελληνικά Γράμματα, 2005.
- [15] Χαραλαμπάκης, Χ., «Λεξικογραφία και ορολογία: Συμπεράσματα από τη σύγκριση δύο σύγχρονων νεοελληνικών λεξικών», *Πρακτικά του 8^{ου} Συνεδρίου Ελληνικής Γλώσσας και Ορολογίας*, 10-12 Νοεμβρίου 2011, Αθήνα: Ελληνικό Επαγγελματικό Επιμελητήριο και Ελληνική Εταιρεία Ορολογίας, σς. 27–39.

Αγγλοελληνικό γλωσσάριο όρων	
Αγγλικός όρος	Ελληνικός όρος
• Electronic excerption	• ηλεκτρονική αποδελίωση
• Electronic slips	• ηλεκτρονικά δελτία
• Collocations	• συμφράσεις
• Literary expressions	• λόγιες εκφράσεις
• Popular expressions	• δημοτικές εκφράσεις
• Fossil words	• απολιθώματα
• Loan words	• δάνειες λέξεις
• Rare words	• σπάνιες λέξεις
• Diminutive words	• υποκοριστικά
• Augmentative words	• μεγεθυντικά
• Nominal adjectives	• ουσιαστικοποιημένα επίθετα
• Neologisms	• νεολογισμοί
• Synecdoches	• συνεκδοχές
• Metaphors	• μεταφορές
• Metonymies	• μετωνυμίες
• Terminology collocations	• συμφράσεις ορολογίας
• Idiomatic phrases	• ιδιωματικές φράσεις
• Support verb constructions	• υποστηρικτικές δομές
• Morphological annotation	• μορφολογικός σχολιασμός
• Annotation tag	• ανάρτημα σχολιασμού

John N. Kazazis

Professor, President of the Centre for the Greek Language

jkazazis@komvos.edu.gr

Rania Voskaki

Researcher at the Centre for the Greek Language

rvoskaki@hotmail.com