

ΠΑΡΑΣΚΕΥΗ Ξ. ΓΙΟΥΛΗ

Υπολογιστική Γλωσσολόγος, PhD

Επίκουρη Καθηγήτρια

Curriculum Vitae

I. ΠΡΟΣΩΠΙΚΑ ΣΤΟΙΧΕΙΑ

Όνομα Παρασκευή
Επίθετο Γιούλη
Πατρώνυμο Ξενοφών
Ημ. Γέννησης 07.07.1969
Τόπος Γέννησης Πάτρα, Αχαΐα
Οικ. Κατάσταση Έγγαμος, 2 παιδιά

Στοιχεία επικοινωνίας

e-mail pgiouli@del.auth.gr; voula@athenarc.gr
Google Scholar <https://scholar.google.com/citations?hl=en&user=dbiF5joAAAAJ>
Orchid Id <https://orcid.org/my-orcid?orcid=0000-0002-1501-9754>

II. ΣΠΟΥΔΕΣ

- 2013 – 2020: **Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Σχολή Οικονομικών και Πολιτικών Επιστημών.** Διδακτορική διατριβή με τίτλο: Σημασιολογικό πεδίο των συναισθημάτων: ταξινόμηση των ρημάτων της Νέας Ελληνικής που δηλώνουν “*ΣΥΝΑΙΣΘΗΜΑ*”. Επιβλέπων καθηγητής: Δρ. Ι.Ε. Σαριδάκης. Μελέτη των συντακτικών και σημασιολογικών ιδιοτήτων των δηλωτικών συναισθήματος ρηματικών κατηγορημάτων της Νέας Ελληνικής με βάση το μεθοδολογικό πλαίσιο του Λεξικού-Γραμματικής. Τα αποτελέσματα της έρευνας αξιοποιήθηκαν για τη δημιουργία γραμματικής με στόχο την αυτόματη επεξεργασία κειμένου και την ανάλυση συναισθήματος σε κείμενο.
- 1995 – 1996: **Πανεπιστήμιο Εδιμβούργου, Μ. Βρετανία.** *MSc in Speech and Language Processing.* MSc thesis: “Topic Chaining and Discourse Structure in Task-oriented dialogues”. Επιβλέπων καθηγητής: Dr Massimo Poesio. Μελέτη και επίλυση της συναναφοράς σε διαλογικά κείμενα. Ξεκινώντας από τη Θεωρία Επικέντρωσης (Centering Theory), η εργασία προτείνει την έννοια της Ολικής Εστίασης (Global Focus) έναντι της Τοπικής Εστίασης (Local Focus), και την αξιοποιεί για τη σύνδεση ενός αναφορικού με το σημείο αναφοράς του, και την επίλυση της Συναναφοράς (Coreference resolution) σε διαλογικά κείμενα.
- 1987 - 1992: **Εθνικό & Καποδιστριακό Πανεπιστήμιο Αθηνών, Φιλοσοφική Σχολή, Τμήμα Φιλολογίας,** Κατεύθυνση Γλωσσολογίας. Πτυχίο Φιλολογίας. Διπλωματική Εργασία: “Η εξέλιξη της Δοτικής”. Επιβλέπουσα καθηγήτρια: Δρ. Δέσποινα Χειλά – Μαρκοπούλου. Στην εργασία αυτή, πραγματοποιήθηκε η βιβλιογραφική επισκόπηση αναφορικά με τη χρήση της δοτικής από την Αρχαία Ελληνική (αττική διάλεκτο) μέχρι την Ελληνιστική κοινή, αλλά και

τη μεσαιωνική Ελληνική. Στο πλαίσιο της εργασίας, και με την αναφορά σε κείμενα των ανωτέρω περιόδων σκιαγραφήθηκε η εξέλιξη της δοτικής πτώσης στην Νέα Ελληνική αλλά και σε διαλέκτους.

1981 – 1987: **Πρότυπος Σχολή Πατρών**, Πάτρα, Αχαΐα.

III. ΕΠΑΓΓΕΛΜΑΤΙΚΗ ΕΜΠΕΙΡΙΑ

2024 – **Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Σχολή Φιλοσοφική, Τμήμα Γερμανικής Γλώσσας και Φιλολογίας**. Επίκουρη Καθηγήτρια, στο γνωστικό αντικείμενο *Γλωσσολογία: Θεωρία και πράξη σε συνάρτηση με την ανάλυση της ελληνικής και της γερμανικής γλώσσας* (ΦΕΚ 1747/14-6-2024 τ. Γ').

1998 – 2024: **Ινστιτούτο Επεξεργασίας του Λόγου, Ερευνητικό Κέντρο “ΑΘΗΝΑ”**. Συνεργαζόμενη ερευνήτρια. Βασική και εφαρμοσμένη έρευνα, σχεδιασμός και την υλοποίηση ερευνητικών και αναπτυξιακών ευρωπαϊκών και εθνικών έργων. Τα ερευνητικά μου ενδιαφέροντα εστιάζουν στη μελέτη και ανάλυση της φυσικής γλώσσας στα επίπεδα της Μορφολογίας, της Σύνταξης και της Σημασιολογίας, και στη δημιουργία ηλεκτρονικών γλωσσικών πόρων (υπολογιστικών λεξικών, σωμάτων κειμένων και γραμματικών), με στόχο την ανάπτυξη και αξιολόγηση υπολογιστικών εργαλείων και εφαρμογών, τόσο της γενικής γλώσσας όσο και ειδικών γλωσσών (ορολογία). Ως βασικό στέλεχος της ομάδας Επεξεργασίας Φυσικής Γλώσσας του ΙΕΛ, συμμετείχα στην ανάπτυξη των βασικών εργαλείων ανάλυσης και επεξεργασίας της Νέας Ελληνικής σε διάφορα επίπεδα γλωσσικής ανάλυσης: Μορφολογία, Σύνταξη, Σημασιολογία. Τα εργαλεία αυτά, μέχρι και σήμερα, αποτελούν τα βασικά αρθρώματα αρχιτεκτονικής για την επεξεργασία της Ελληνικής. Στο πλαίσιο αυτό, έχω ασχοληθεί με την κατάρτιση και επισημείωση ηλεκτρονικών σωμάτων κειμένων τα οποία επιτρέπουν τη μελέτη γλωσσικών φαινομένων της γενικής γλώσσας, αλλά και ειδικών γλωσσών και γλωσσικών ποικιλιών. Κομβική θέση στην έρευνά μου κατέχουν ζητήματα αναπαράστασης της γλωσσικής πληροφορίας σε λεξικά τόσο για ανθρώπους-χρήστες (π.χ. λεξικά της Ελληνικής για ξένους, για μετανάστες, ερμηνευτικό λεξικό για το Γυμνάσιο, μικρό λεξικό του θρακικού ιδιώματος, κ.λπ.) όσο και για υπολογιστικά συστήματα με στόχο την υποστήριξη της διάδρασης ανθρώπου-μηχανής (π.χ. εννοιολογικό λεξικό της Νέας Ελληνικής, λεξικό πολυλεκτικών εκφράσεων, λεξικό FrameNet, ειδικά ορολογικά λεξικά, κ.λπ.). Τέλος, έχω ασχοληθεί με την ανάπτυξη κειμενικών συλλογών (digital libraries) που αποτυπώνουν την ιστορική εξέλιξη της Νέας Ελληνικής αλλά και γεωγραφικές της ποικιλίες, και την προσαρμογή των εργαλείων επεξεργασίας της Ελληνικής με στόχο την ανάλυση και δεικτοδότηση των κειμένων που συγκροτούν τις εν λόγω συλλογές. Τα τελευταία χρόνια, η έρευνά μου επικεντρώνεται ιδιαιτέρως στη

διεπαφή σύνταξης – σημασιολογίας, στη μελέτη των πολυλεκτικών εκφράσεων, καθώς αυτές αποτελούν τροχοπέδη σε πολλές εφαρμογές και γλωσσικές τεχνολογίες, και στην ανάλυση συναισθήματος και συναισθηματικού λόγου.

1996 – 1998: **Εργαστήριο Ενσύρματης Τηλεπικοινωνίας, Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών, Πανεπιστήμιο Πατρών.** Βασική και εφαρμοσμένη έρευνα με στόχο τη μορφολογική ανάλυση της Ελληνικής (σύστημα μορφολογικής ανάλυσης που ενσωματώθηκε σε λεξικά), αλλά και συμμετοχή στον σχεδιασμό και ανάπτυξη λεξικών της Νέας Ελληνικής (Λεξικό της Νέας Ελληνικής για Άγγλους χρήστες, δίγλωσσο ελληνο-αγγλικό και αγγλο-ελληνικό λεξικό ΚΟΡΑΗΣ).

1993 – 1995 **Εργαστήριο Ενσύρματης Τηλεπικοινωνίας, Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών, Πανεπιστήμιο Πατρών.** Συμμετοχή στο λεξικογραφικό έργο του εργαστηρίου και την ανάπτυξη δύο ηλεκτρονικών λεξικών προφοράς και χρήσης: (α) της Νέας Ελληνικής για ξένους, και (β) της ελληνικής διαλέκτου της Νότιας Ιταλίας (γραιοκαννική/γκρίκο). Υπήρξα μέλος της λεξικογραφικής ομάδας για την ανάπτυξη του πλαισίου (frame) για το δίγλωσσο Ελληνο-αγγλικό Λεξικό (σε συνεργασία με το λεξικογραφικό κέντρο Cobuild και τον εκδοτικό οίκο Collins). Τέλος, συμμετείχα στις εργασίες για τη δημιουργία - κατά το πρότυπο του Bank of English - του Σώματος Κειμένων (ΣΚ) της ελληνικής γλώσσας (60.000.000 λέξεις) του εργαστηρίου, καθώς επίσης και του περιβάλλοντος διαχείρισης και επεξεργασίας Σωμάτων Κειμένων ΛΕΞΙΘΗΡΑΣ.

IV. ΔΙΔΑΚΤΙΚΗ ΕΜΠΕΙΡΙΑ

2024 – 2025 Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Σχολή Φιλοσοφική, Τμήμα Γερμανικής Γλώσσας και Φιλολογίας. Λεξικολογία και Λεξικογραφία, Υπολογιστική Λεξικογραφία, Υπολογιστική Γλωσσολογία, Ψηφιακές Ανθρωπιστικές Επιστήμες.

2023 Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Σχολή Οικονομικών και Πολιτικών Επιστημών, Τμήμα Τουρκικών και Σύγχρονων Ασιατικών Σπουδών. Εντεταλμένη διδάσκουσα. Αυτοδύναμη διδασκαλία. **Ηλεκτρονική Λεξικογραφία και Ορογραφία (70255).**

2016 – 2018 Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών. Σχολή Οικονομικών και Πολιτικών Επιστημών. Τμήμα Τουρκικών και Σύγχρονων Ασιατικών Σπουδών. Μεταπτυχιακό πρόγραμμα σπουδών στη Μετάφραση και Διερμηνεία. Μάθημα: **Εφαρμοσμένη Ορογραφία (TURKMAS320).**

2012 – 2014	University of Dundee, Athens Campus. MA (Hons) in Education. Module 1GR1: Greek Linguistics and Associated Pedagogy.
2005 – 2006	Athens Metropolitan College. MSc in Translation Studies. Γνωστικό αντικείμενο: <i>Computer Assisted Translation Tools</i> .
1993-1996:	Koutsandonis School of Languages, Πάτρα, Καθηγήτρια Αγγλικής Γλώσσας.

V. ΕΡΕΥΝΑ: ΣΥΜΜΕΤΟΧΗ ΣΕ ΕΡΕΥΝΗΤΙΚΑ ΚΑΙ ΑΝΑΠΤΥΞΙΑΚΑ ΕΡΓΑ

2023 – 2027. COST Action CA22126 - European Network On Lexical Innovation (ENEOLI). Αντικείμενο του έργου είναι η διεξαγωγή πρωτογενούς έρευνας στο πεδίο της νεολογίας. Ειδικότερα, το έργο στοχεύει στα ακόλουθα: (i) καθορισμός της ορολογίας που άπτεται του πεδίου της νεολογίας, και δημιουργία ηλεκτρονικού (digitally born) ειδικού πολύγλωσσου σχετικού γλωσσαρίου (το οποίο αυτή τη στιγμή δεν υπάρχει), με σκοπό τη διευκόλυνση της έρευνας σε διεθνές επίπεδο, (ii) προσαρμογή υπολογιστικών μεθόδων και εργαλείων για τον εντοπισμό νεολογισμών, (iii) η συγκριτική μελέτη του φαινομένου της νεολογίας στις ευρωπαϊκές γλώσσες, με ειδική έμφαση στο φαινόμενο του δανεισμού, και (iv) η εκπαίδευση μεταφραστών, υπολογιστικών γλωσσολόγων, δημοσιογράφων, συγγραφέων τεχνικών και ειδικευμένων κειμένων αλλά και εκπαιδευτικών, μέσω ενός συγκεκριμένου πρωτοκόλλου που θα μπορούσε να αναπαραχθεί για οποιαδήποτε ευρωπαϊκή γλώσσα. Τέλος, το έργο στοχεύει στην ενημέρωση του ευρύτερου κοινού (των μη ειδικών) αναφορικά με το φαινόμενο της νεολογίας, με τη εμπλοκή στο έργο θεσμικών φορέων αλλά και του ευρύτερου κοινού, και την συνακόλουθη ανάδειξη των κοινωνικών προεκτάσεων του φαινομένου, σε δύο άξονες: προώθηση της δημιουργικότητας στη μητρική γλώσσα, σαφήνεια στην επικοινωνία αλλά και στην επιστήμη.

2022 – 2026. COST Action CA21167 - Universality, diversity and idiosyncrasy in language technology (UniDive). Αντικείμενο του έργου είναι η μελέτη των γλωσσικών καθολικών, της γλωσσικής ποικιλότητας και των ιδιοσυγκρασιακών στοιχείων στη γλώσσα, σε συνάρτηση με τη γλωσσική τεχνολογία και την επεξεργασία φυσικής γλώσσας (NLP). Η πρόσβαση στις συνεχώς αυξανόμενες ποσότητες δεδομένων (δομημένων ή μη), ιδίως γλωσσικών δεδομένων, και η εξαγωγή γνώσης από αυτά με τη βοήθεια τεχνικών και αλγορίθμων της επιστήμης των δεδομένων (data science), στηρίζεται στην ύπαρξη καθολικών στη γλώσσα. Ταυτόχρονα, όμως, προϋποθέτει την επίλυση προκλήσεων που θέτει η κάθε μορφής γλωσσική ποικιλότητα. Στο επίπεδο της θεωρίας, η έρευνα αποσκοπεί στην καλύτερη κατανόηση των καθολικών στη γλώσσα, και την ομογενοποίηση της σχετικής ορολογίας, ενώ από πρακτικής απόψεως, το έργο αποσκοπεί στη δημιουργία γλωσσικών πόρων και εργαλείων που καλύπτουν, με ενιαίο τρόπο, μεγαλύτερη ποικιλία γλωσσικών φαινομένων σε μεγάλο αριθμό γλωσσών, συμπεριλαμβανομένων εκείνων για τις οποίες δεν υπάρχουν πολλοί γλωσσικοί πόροι (low-resourced languages) αλλά και εκείνων που κινδυνεύουν με εξαφάνιση. Συμμετοχή στο έργο ως μέλος της Επιτροπής Διαχείρισης (Management Committee) αλλά και

ως υπεύθυνη της ομάδας εργασίας 2 (Working Group 2) Διεπαφή σωμάτων κειμένων – λεξικού (Corpus – Lexicon Interface).

2021 - 2022. Δράση ΚΑΛΛΙΠΟΣ+. Ανοικτά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα. Πρόσκληση-VI1. Ανθρωπιστικές Επιστήμες και Τέχνες. Συγγραφή νέου διδακτικού εγχειριδίου για προπτυχιακά και μεταπτυχιακά μαθήματα. Τίτλος εγχειριδίου: Φωτοπούλου, Αγγελική και Γιούλη, Παρασκευή. [Υπό έκδοση]. *Περί Λεξιλογίου. Λεξικο-λογία και Λεξικο-γραφία: έντυπη, ηλεκτρονική, υπολογιστική.* Εκδόσεις ΚΑΛΛΙΠΟΣ. Συμμετοχή στη συγγραφή ως συν-συγγραφέας, υπεύθυνη για το 50% του περιεχομένου.

2021 - . Μυθολογικές Διαδρομές στην Ανατολική Μακεδονία – Θράκη (ΟΠΣ (MIS) 5047101). Στόχος του έργου είναι η δημιουργία ψηφιακής υποδομής, πληροφοριακού ιστότοπου και εφαρμογής για κινητά που θα επιτρέπουν τη δημιουργία προσωποποιημένων διαδρομών σε περιοχές της περιφέρειας με απώτερο στόχο την υποστήριξη της ολοκληρωμένης τουριστικής εμπειρίας στην περιοχή της Ανατολικής Μακεδονίας και Θράκης με άξονα τους *μύθους*, τη *γλώσσα* και τον *πολιτισμό*. Στόχος του έργου είναι η καταγραφή των μύθων (όπως αυτοί αναδύονται μέσα από κείμενα της Αρχαίας Ελληνικής και Λατινικής Λογοτεχνίας), ο σχετικός σχολιασμός και η διασύνδεσή τους με το σύγχρονο τοπίο της περιφέρειας (φυσικό, αρχιτεκτονικό, αρχαιολογικό, πολιτιστικό, γλωσσικό) και η σύνδεσή τους με ψηφιακούς χάρτες. Στο πλαίσιο του έργου θα αναπτυχθεί ειδική εκπαιδευτική εφαρμογή για την παιδαγωγική αξιοποίηση του παραγόμενου ερευνητικού υλικού από διδάσκοντες και διδασκόμενους στη Δευτεροβάθμια και Τριτοβάθμια εκπαίδευση. Η Πράξη θα ενισχύσει ουσιαστικά την τρέχουσα Εθνική Υποδομή για τις Ψηφιακές Ανθρωπιστικές Επιστήμες και για την Ελληνική Γλωσσική Έρευνα και Καινοτομία ΑΠΟΛΛΩΝΙΣ. Ρόλος μου στο έργο είναι (α) η κατάρτιση (πολυτροπικού) σώματος κειμένων από τα πρωτογενή δεδομένα (κείμενα αρχαίας ελληνικής και λατινικής γραμματείας, περιηγητικά, τουριστικά, κ.α. κείμενα), (β) η δεικτοδότηση και επεξεργασία (τμήματος) του εν λόγω σώματος κειμένων με στόχο την σύνταξη γλωσσικού θησαυρού ο οποίος θα αποτελέσει τη ραχοκοκαλιά της υποδομής. Ο γλωσσικός θησαυρός αποτελείται από όρους σχετικούς με ειδικά θεματικά πεδία της εφαρμογής (γαστρονομία, γεωγραφία, αρχαιολογία, λαογραφία, χλωρίδα, πανίδα, κ.λπ.), αλλά και κύρια ονόματα (τοπωνύμια, ονόματα προσώπων, μυθολογικών προσώπων, κ.λπ.) καλύπτοντας ένα ευρύ πλαίσιο στους άξονες του χώρου και του χρόνου, αλλά και δύο γλώσσες (Ελληνική και Αγγλική) με στόχο (α) την δεικτοδότηση και ανάκτηση των κειμένων, και (β) την υποβοήθηση της μεταφραστικής διαδικασίας.

2019 - 2022. ELEXIS – European Lexicographic Infrastructure. Το έργο αποσκοπεί στην προώθηση της συνεργασίας μεταξύ όλων των εμπλεκόμενων στη Λεξικογραφία και τη λεξικογραφική πρακτική (ερευνητική κοινότητα, πανεπιστήμια, εκδοτικοί οίκοι) σε παν-ευρωπαϊκό επίπεδο. Απώτερος στόχος είναι η δημιουργία κοινά αποδεκτών προτύπων (standards), εργαλείων και πρακτικών για την δημιουργία, αξιοποίηση, συντήρηση και επαναχρησιμοποίηση λεξικογραφικών πόρων. Συμμετοχή στο έργο ως εκπρόσωπος του ΙΕΛ/ΑΘΗΝΑ, με στόχο την εναρμόνιση επιλεγμένων λεξικογραφικών πόρων του ΙΕΛ με τα

διεθνή πρότυπα, και τη διασύνδεσή τους με άλλους λεξικούς πόρους εκτός του οργανισμού στο πρότυπο των γλωσσικών διασυνδεδεμένων δεδομένων (Linked Linguistic Data).

2019 - . The Global FrameNet: Shared Annotation Task. Αντικείμενο του έργου είναι η δημιουργία και στοίχιση γλωσσικών πόρων οι οποίοι αναπτύσσονται σύμφωνα με τις αρχές της Σημασιολογίας Πλαισίων (Frame Semantics). Στην κατεύθυνση αυτή, η επισημείωση παράλληλων ή συγκρίσιμων κειμένων σε διαφορετικές γλώσσες, έχει ως απώτερο στόχο στον εντοπισμό και την αξιολόγηση των αποκλίσεων και διαφορών που (πρέπει να) υπάρχουν μεταξύ των εν λόγω λεξικών πόρων (FrameNets) για διαφορετικές γλώσσες, και τον υπολογισμό της πολυπλοκότητας που ενέχει η στοίχιση των πόρων αυτών. Απώτερο στόχο του έργου, αποτελεί η δημιουργία συλλογής κειμένων σχολιασμένων με πλαίσια για πολλές γλώσσες, τα οποία θα μπορούν να χρησιμοποιηθούν στο μέλλον, ως δεδομένα εκπαίδευσης και αξιολόγησης σε διάφορες πολύ-γλωσσικές εφαρμογές. Το έργο τελεί υπό την επίβλεψη του καθ. Colin Baker και της ομάδας του International Computer Science Institute του πανεπιστημίου του Berkeley, ΗΠΑ (<https://www.globalframenet.org/annotation>). Συμμετοχή στο έργο ως επιστημονική υπεύθυνη της ομάδας του Ελληνικού FrameNet.

2017 – 2021. ΑΠΟΛΛΩΝΙΣ: Εθνική Υποδομή για τις Ψηφιακές Ανθρωπιστικές Τέχνες και Επιστήμες και για την Γλωσσική Έρευνα και Καινοτομία. ΕΠΑνΕΚ / Πρόσκληση 039 "Ενίσχυση των Υποδομών Έρευνας και Καινοτομίας". Εθνική κατανεμημένη υποδομή που προέκυψε από τη συγχώνευση των υποδομών: CLARIN και DARIAH. Η συνένωση των δύο υποδομών θα παράσχει στην ερευνητική κοινότητα, αλλά και στη βιομηχανία και το ευρύ κοινό, πρόσβαση σε αξιόλογες συλλογές διαλειτουργικών γλωσσικών και πολιτιστικών πόρων και σε καινοτόμες υπηρεσίες επεξεργασίας αυτών, διευκολύνοντας και προωθώντας την ψηφιακή διάσταση των επιστημών. Συμμετοχή στις ενέργειες τεκμηρίωσης γλωσσικών πόρων και εργαλείων του ΙΕΛ/ΑΘΗΝΑ (curator).

2017 – 2019. Υπολογιστικές Επιστήμες και Τεχνολογίες Δεδομένων, Περιεχομένου και Αλληλεπίδρασης (ΟΠΣ 5002437). Συμμετοχή στην υλοποίηση του υποέργου 1 «Γλωσσικές Τεχνολογίες Ανάλυσης Περιεχομένου και Αλληλεπίδρασης», και πιο συγκεκριμένα, στην Ενότητα Εργασίας «ΕΕ5: Υποδομή Γλωσσικών Πόρων», η οποία περιλαμβάνει δράσεις που αφορούν (i) στην ανάπτυξη κειμενικών και λεξικών γλωσσικών πόρων, (ii) στον εμπλουτισμό υπάρχοντων γλωσσικών πόρων, και, εν τέλει, (iii) στη διασύνδεση και αποθήκευσή τους σε υφιστάμενα αποθετήρια και υποδομές. Συμμετοχή στο έργο στο πλαίσιο της έρευνας για τον εμπλουτισμό και τη σύνταξη του ηλεκτρονικού εννοιολογικού λεξικού της Ελληνικής, τη μεταφορά του σε API, αλλά και τη διασύνδεσή του με υφιστάμενα λεξικά του ΙΕΛ στο πρότυπο των γλωσσικών διασυνδεδεμένων δεδομένων (Linked Linguistic Data). Στο πλαίσιο αυτό, έχω εργαστεί για τη διασύνδεση του εννοιολογικού λεξικού (α) με το λεξικό της Ελληνικής Νοηματικής Γλώσσας, και (β) με το ελληνικό μέρος του σημασιολογικού λεξικού SIMPLE.

2014 – 2018. TextLink: Structuring Discourse in Multilingual Europe (ISCH COST Action IS1312). Στόχος του έργου ήταν η δημιουργία υποδομής γλωσσικών πόρων (λεξικών, σωμάτων κειμένων) και εργαλείων για τη δια-γλωσσική μελέτη του τρόπου οργάνωσης του λόγου (discourse structure). Ειδικότερα, το έργο προέβλεπε τις ακόλουθες δράσεις: (1) εντοπισμός σχετικών γλωσσικών πόρων (discourse lexica and corpora, εργαλείων σχολιασμού) και δημιουργία μιας πύλης μέσω της οποίας θα διευκολύνεται η πρόσβαση σε αυτούς, (2) ορισμός βασικών αρχών και απαιτήσεων ενός σχήματος σχολιασμού, (3) οργάνωση των αρχών αυτών σε ένα κοινό ταξινομικό σχήμα, (4) προώθηση του ταξινομικού σχήματος με στόχο την υιοθέτησή του σε μελλοντικές απόπειρες σχολιασμού κειμένων, και (5) προώθηση της πύλης, των πόρων και του ταξινομικού σχήματος στην επιστημονική κοινότητα. Τελικώς, το έργο TextLink αποσκοπούσε στην ομογενοποίηση των μέχρι σήμερα διάσπαρτων γλωσσικών πόρων, και στην διευκόλυνση της προσβασιμότητας σε αυτούς, με στόχο την βέλτιστη αξιοποίησή τους στη μετάφραση, τη λεξικογραφία, τη γλωσσική τεχνολογία, τη διδασκαλία, κλπ. Στο πλαίσιο του έργου ασχολήθηκαν με τις προδιαγραφές των λεξικών και με τις προδιαγραφές του σχήματος σχολιασμού, ενώ ξεκίνησε ο σχολιασμός ελληνικού σώματος κειμένων στο επίπεδο αυτό.

2013 – 2017. ENeL: European Network of electronic Lexicography. (COST Action IS1305). Η Πράξη αποσκοπούσε στη δημιουργία ενός Ευρωπαϊκού δικτύου επιστημόνων που ασχολούνται με τη λεξικογραφία, με στόχο τη συστηματική και αποτελεσματική αντιμετώπιση των προκλήσεων που θέτουν τα ηλεκτρονικά μέσα και το διαδίκτυο στον τρόπο ανάπτυξης, διάθεσης και χρήσης των λεξικών. Ειδικότερα, νέες τεχνολογίες θέτουν προκλήσεις αυτές, όπως: (α) παροχή στους χρήστες πρόσβασης σε ακαδημαϊκά λεξικά κατά τρόπον ώστε να γεφυρώνεται το χάσμα που υπάρχει σήμερα ανάμεσα στο ευρύ κοινό και τα λεξικά αυτά, (β) ανταλλαγή τεχνογνωσίας, υιοθέτηση κοινών προτύπων (common standards) και λύσεων σε ευρεία κλίμακα και με τρόπο συστηματικό, καθώς και (γ) υιοθέτηση κοινών προσεγγίσεων στην ηλεκτρονική λεξικογραφία (e-lexicography) που συνιστούν τη βάση ενός νέου είδους λεξικογραφικής πρακτικής, η οποία θα αντανakλά και θα προβάλλει πλήρως την κοινή κληρονομιά του λεξιλογίου των ευρωπαϊκών γλωσσών. Στο πλαίσιο του έργου, συμμετείχα στον σχεδιασμό και την υλοποίηση πανευρωπαϊκής έρευνας σχετικά με τη θέση των λεξικών και της λεξικογραφίας στα κράτη-μέλη της Ευρωπαϊκής Ένωσης, ως υπεύθυνη του σχετικού ερωτηματολογίου για την Ελλάδα. Τα αποτελέσματα της έρευνας δημοσιεύθηκαν σε δύο άρθρα σε επιστημονικά περιοδικά.

2013 – 2017. PARSEME: PARSing and Multi-word Expressions. Towards linguistic precision and computational efficiency in natural language processing. (COST Action IC1207). Αντικείμενο του έργου υπήρξε η μοντελοποίηση και επεξεργασία των Πολυλεκτικών Εκφράσεων (Multiword Expressions), δηλαδή ακολουθιών λέξεων με απρόβλεπτες ιδιότητες όπως "το παίρνω πάνω μου" ή "νίπτω τας χείρας μου", οι οποίες δημιουργούν προβλήματα σε διάφορες εφαρμογές. Στο πλαίσιο του έργου μελετήθηκαν 14 Ευρωπαϊκές γλώσσες – μεταξύ των οποίων και η Ελληνική. Στα παραδοτέα περιλαμβάνονται αναβαθμισμένοι γλωσσικοί πόροι και εργαλεία καθώς και οδηγίες για βέλτιστες πρακτικές όσον αφορά

γλωσσικά μοντέλα αιχμής που λαμβάνουν υπόψιν τους τις Πολυλεκτικές Εκφράσεις, ενώ αποτέλεσμα του έργου είναι και η διοργάνωση από το 2017 και εξής, της σειράς των διεθνών διαγωνισμών συστημάτων αναγνώρισης ρηματικών πολυλεκτικών εκφράσεων (Shared Tasks on Verbal Multiword Expressions) και του σχετικού θεματικού συνεδρίου. Στο πλαίσιο αυτό, ασχολήθηκα με τη μοντελοποίηση των πολυλεκτικών εκφράσεων της Ελληνικής, ενώ υπήρξα συν-διοργανώτρια των διαγωνισμών (Shared Tasks) συμμετέχοντας στη συγγραφή των σχετικών προδιαγραφών επισημείωσης, αλλά και ως υπεύθυνη της ελληνικής ομάδας επισημείωσης (Language Leader), αλλά και υπεύθυνη για τον συντονισμό ομάδων άλλων γλωσσών (Language Group Leader).

2015 – 2017. ΑΝΑΛΥΤΙΚΗ ΠΕΡΙΕΧΟΜΕΝΟΥ ΚΑΙ ΠΟΛΥΤΡΟΠΙΚΗ ΕΠΙΚΟΙΝΩΝΙΑ – ΠΟΛΥΤΡΟΠΟΝ. (Αναπτυξιακές Προτάσεις Ερευνητικών Φορέων – ΚΡΗΠΙΣ, Υπουργείο Παιδείας, Διά Βίου Μάθησης & Θρησκευμάτων – Γ.Γ.Ε.Τ.). Άξονας 3: Αναλυτική γραπτού λόγου στη Διαχείριση Γνώσης. Στο πλαίσιο του έργου ασχολήθηκα με τη σχεδίαση και ανάπτυξη του εννοιολογικού λεξικού του ΙΕΛ, το οποίο θα αξιοποιηθεί οριζόντια σε διάφορες εφαρμογές και γλωσσικές τεχνολογίες.

2012 – 2015. CLARIN-EL. ΕΠΑΝ ΙΙ / Ανταγωνιστικότητα - Επιχειρηματικότητα. Το CLARIN-EL είναι το ελληνικό σκέλος του έργου CLARIN, μιας πανευρωπαϊκής προσπάθειας να συγκεντρωθούν και να διατεθούν στην ερευνητική κοινότητα μέσω μιας διαδικτυακής Ερευνητικής Υποδομής γλωσσικοί πόροι (κείμενα, καταγραφές προφορικού λόγου, πολυμεσικά δεδομένα, λεξικά, ορολογικά γλωσσάρια κτλ.) σε όλες τις γλώσσες, και εργαλεία γλωσσικής τεχνολογίας για την επεξεργασία τους (μορφολογικοί / συντακτικοί αναλυτές, εργαλεία επισημείωσης, στατιστικά εργαλεία κτλ.). Στόχος του έργου είναι (α) να διαθέσει στην ερευνητική κοινότητα την υποδομή για τη συγκέντρωση και την τεκμηρίωση γλωσσικών πόρων και εργαλείων γλωσσικής τεχνολογίας για τα Ελληνικά σε ένα δίκτυο αποθετηρίων τα οποία θα εμπλουτίζονται διαρκώς από τα μέλη του δικτύου, (β) να παρέχει διαδικτυακές υπηρεσίες πρόσβασης σε πόρους και εργαλεία, (γ) να διερευνήσει το νομικό καθεστώς πρόσβασης σε αυτά, και να παρέχει νομικές συμβουλευτικές υπηρεσίες σε παρόχους και χρήστες, και (γ) να ενημερώσει και να εκπαιδεύσει ερευνητές και ευρύ κοινό για τις δυνατότητες στη χρήση των γλωσσικών πόρων και εργαλείων. Συμμετοχή στο έργο με τη τεκμηρίωση γλωσσικών πόρων και εργαλείων του ΙΕΛ/ΑΘΗΝΑ.

2012 – 2015. LangTERRA: Enhancing the Research Potential of ILSP/"Athena" R.C. in Language Technology in the European Research ERA. Το LangTERRA αποτελεί ένα έργο αριστείας που αντιπροσωπεύει την ισχυρή δέσμευση του ΙΕΛ για τη συνεχή ενίσχυση των ικανοτήτων του. Για τη διαμόρφωση του στρατηγικού ερευνητικού του σχεδίου του ΙΕΛ στο διαρκώς εξελισσόμενο πεδίο της Γλωσσικής Τεχνολογίας, έχουν επιλεγεί τέσσερα ερευνητικά πεδία προτεραιότητας που αναμένεται να βρεθούν στην αιχμή της γλωσσικής τεχνολογίας τα επόμενα χρόνια: (i) Πολυτροπική επικοινωνία, (ii) Επεξεργασία πολυμέσων, (iii) Γλωσσομάθεια και μαθησιακές δυσκολίες, (iv) Εξόρυξη κειμένου. Στο πλαίσιο των παραπάνω αξόνων και με στόχο την ενίσχυση των ικανοτήτων και της ανταγωνιστικότητας του ΙΕΛ, το

LangTERRA θα αξιοποιήσει μια σειρά εργαλείων και θα στηρίξει δραστηριότητες κινητοποίησης των πόρων και υποδομών του ΙΕΛ. Πρόκειται για τη δικτύωση με 14 ερευνητικούς φορείς παγκόσμιας κλάσης από 9 διαφορετικές ευρωπαϊκές χώρες. Επιπλέον, περιλαμβάνει την πρόσληψη πεπειραμένων ερευνητών στα ερευνητικά πεδία προτεραιότητας, τη διοργάνωση εκδηλώσεων δικτύωσης και διάχυσης των ερευνητικών αποτελεσμάτων, όπως συνέδρια και ημερίδες, και την απόκτηση state-of-the-art ερευνητικού εξοπλισμού.

2011 – 2013. Ε.Π. ΨΗΦΙΑΚΗ ΣΥΓΚΛΙΣΗ (ΕΣΠΑ 2007-2013): e-Milang. Βελτιωμένες Ψηφιακές Υπηρεσίες Στοχευμένης Πληροφόρησης Μεταναστών. Η Πράξη *e-MiLang* είχε ως αντικείμενο την συγκέντρωση, ανάπτυξη, ψηφιοποίηση και οργάνωση πληροφοριακού υλικού, προσανατολισμένου στην ικανοποίηση των αναγκών διαβίωσης των πολυπληθέστερων μεταναστευτικών ομάδων στην Ελλάδα, καθώς και την διάχυση του υλικού αυτού με ηλεκτρονικά μέσα, με σκοπό την υποστήριξη της διαδικασίας ομαλής προσαρμογής και ένταξης των στοχευόμενων ομάδων μεταναστών στην ελληνική κοινωνία και πολιτεία. Το πληροφοριακό χαρακτήρα ψηφιακό περιεχόμενο που αναπτύχθηκε στα πλαίσια του έργου αποτελείται από: (α) 9 δίγλωσσα λεξικά (από τα Ελληνικά προς καθεμία από τις γλώσσες Αλβανικά, Βουλγαρικά, Ρουμανικά, Ρωσικά, Σερβικά, Πολωνικά, Κινέζικα-Μανδαρινικά, Αραβικά και Αγγλικά), (β) 9 δίγλωσσα (Ελληνικά και καθεμία από τις γλώσσες υποστήριξης) τεύχη-οδηγούς, που θα ενημερώνουν και θα καθοδηγούν τους μετανάστες στις συναλλαγές τους με το Δημόσιο, και (γ) 9 δίγλωσσα τεύχη (Ελληνικά και καθεμία από τις γλώσσες υποστήριξης) με πληροφοριακό υλικό, που θα διευκολύνει τους μετανάστες στην εξοικείωσή τους με το γενικότερο πολιτισμικό και κοινωνικό πλαίσιο της Ελλάδας. Στο πλαίσιο του έργου, ασχολήθηκαν με όλες τις εργασίες κατάρτισης και γλωσσικής επεξεργασίας των κειμενικών δεδομένων, την συγγραφή των προδιαγραφών των λεξικών, αλλά και με την κατάρτιση του λημματολογίου και τη σύνταξη του ελληνικού τμήματος των λεξικών.

2012. Διαδικτυακή εκπαιδευτική πύλη ψυχαγωγίας και μάθησης για μικρά παιδιά (Ε.Π. ΨΗΦΙΑΚΗ ΣΥΓΚΛΙΣΗ, ΕΣΠΑ 2007-2013). Αντικείμενο του έργου υπήρξε η δημιουργία ενός εκπαιδευτικού web portal, ειδικά σχεδιασμένου για μικρά παιδιά 3-7 ετών μέσα στο οποίο θα παίζουν και παράλληλα θα μαθαίνουν. Στόχος της πράξης ήταν η ολόπλευρη ανάπτυξη της προσωπικότητας των μικρών παιδιών μέσα από την ανάπτυξη της γλωσσικής τους έκφρασης, της μουσικής τους παιδείας. Παράλληλα δίνεται η δυνατότητα για έγκαιρη ανίχνευση και αντιμετώπιση των τυχόν προβλημάτων λόγου και ομιλίας, ώστε να καταστεί ομαλή η σχολική τους ένταξη. Εργάστηκα στο τελευταίο στάδιο του έργου για τον έλεγχο, την πιλοτική χρήση και αξιολόγηση του εκπαιδευτικού πακέτου.

2012 – 2014. Analysis and Evaluation of Comparable Corpora for Under-Resourced Areas of Machine Translation (ACCURAT) - FP7 – ICT. Στόχος του έργου ήταν να διερευνήσει και να αξιολογήσει νέες μεθόδους και τεχνικές, οι οποίες επιτρέπουν την αξιοποίηση των συγκρίσιμων σωμάτων κειμένων στη ΜΜ, με στόχο την αποτελεσματική αντιμετώπιση βελτίωση της ποιότητας της μετάφρασης ιδίως για γλώσσες ή/και ζεύγη γλωσσών με ελάχιστους διαθέσιμους πόρους και για εξειδικευμένα γνωστικά πεδία.

Το έργο ανέπτυξε ένα πλήρως λειτουργικό μοντέλο, το οποίο υποστηρίζει: (α) την δυνατότητα αυτόματης ανάκτησης συγκρίσιμων σωμάτων κειμένων από το Διαδίκτυο και από άλλες πηγές, (β) την ανάλυση και αξιολόγηση των συγκρίσιμων σωμάτων κειμένων, με βάση μετρικές συγκρισιμότητας, (γ) την στοίχιση συγκρίσιμων σωμάτων κειμένων (στο επίπεδο του κειμένου, της πρότασης, της φράσης και της λέξης) και την εξαγωγή λεξιλογικών, ορολογικών και άλλων γλωσσικών δεδομένων από αυτά, (δ) ενσωμάτωση των στοιχισμένων κειμένων και των στοιχισμένων γλωσσικών δεδομένων σε υπάρχοντα συστήματα MM (στατιστικά ή/και κανονιστικά), και (ε) ποσοτικοποιημένη εκτίμηση της βελτίωσης που προκύπτει από την αξιοποίηση της ανωτέρω μεθόδου στα υφιστάμενα συστήματα MM. Συμμετείχα σε όλα τα στάδια του έργου, από τον αρχικό σχεδιασμό και την υποβολή της πρότασης, μέχρι και την υλοποίηση.

2008 – 2009.ΛΟΓΟΤΕΧΝΙΚΑ ΠΑΡΑΛΛΗΛΑ. Διαδικτυακός Τόπος και CD-ROM για την Μελέτη και Προβολή της Πολιτιστικής Κληρονομιάς των Διασυνοριακών Περιοχών Ελλάδας – Βουλγαρίας. (INTERREG-PHARE CBC – ΕΛΛΑΔΑ – ΒΟΥΛΓΑΡΙΑ). Μέσα από το Έργο «Λογοτεχνικά Παράλληλα» αναπτύχθηκαν σύγχρονες και τεχνολογικά προηγμένες πλατφόρμες, οι οποίες παρέχουν διευρυμένες δυνατότητες για την μελέτη και ανάδειξη της πολιτιστικής κληρονομιάς της διασυνοριακής περιοχής Ελλάδας – Βουλγαρίας (περιφέρεια Θράκης και όμορες Βουλγαρικές επαρχίες Smolyan, Blagoevgrad, Kardjali, Khaskovo), με αφετηρία την λογοτεχνία, τη λαογραφία και τη γλώσσα των δύο λαών. Για τον σκοπό αυτό, καταρτίστηκαν δύο ηλεκτρονικά σώματα κειμένων στην Ελληνική και Βουλγαρική, καθένα εκ των οποίων είναι αντιπροσωπευτικό της λογοτεχνικής και λαογραφικής παραγωγής των όμορων περιοχών. Τα σώματα κειμένων καλύπτουν τις εν λόγω γεωγραφικές περιοχές σε συγκεκριμένο χρονικό πλαίσιο (19^{ος} – 20^{ος} αιώνας), και επομένως αντικατοπτρίζουν τη γλώσσα στον συγκεκριμένο χώρο και χρόνο. Για την καλύτερη και πιο αποτελεσματική διαχείριση των σωμάτων κειμένων, αλλά και για τη διευκόλυνση των αναζητήσεων μέσα σε αυτά, κρίθηκε αναγκαία η γλωσσική επεξεργασία τους σε διάφορα επίπεδα γλωσσικής ανάλυσης (μορφοσύνταξη, λημματοποίηση, αναγνώριση ονοματικών οντοτήτων, αναγνώριση όρων στα λαογραφικά κείμενα). Η επεξεργασία αυτή πραγματοποιήθηκε με τρόπο ημι-αυτόματο, δηλαδή με τη χρήση υφιστάμενων εργαλείων επεξεργασίας φυσικής γλώσσας για την Ελληνική και τη Βουλγαρική, και τον χειροκίνητο έλεγχο και διόρθωση του εξαγομένου από γλωσσολόγους. Για τις ανάγκες του έργου, υφιστάμενα εργαλεία επεξεργασίας φυσικής γλώσσας προσαρμόστηκαν ώστε να μπορούν να επεξεργαστούν τα εν λόγω κειμενικά δεδομένα. Επίσης, για την καλύτερη προσπέλαση του κειμενικού υλικού, δημιουργήθηκαν ειδικά (δίγλωσσα) λεξικά για ανθρώπους – χρήστες. Τόσο τα σώματα κειμένων όσο και τα λεξικά είναι προσβάσιμα τόσο στο ευρύτερο όσο και στο ειδικό κοινό μέσω ειδικού διαδικτυακού τόπου αλλά και CD-ROM, για επιστημονική, εκπαιδευτική και πολιτιστική χρήση. Υπήρξα αναπληρώτρια επιστημονική υπεύθυνη του έργου, υπεύθυνη για την σύνταξη των προδιαγραφών κατάρτισης και επεξεργασίας του ηλεκτρονικού σώματος κειμένων, για την κατάρτιση των γλωσσικών πόρων της Ελληνικής, αλλά και για την προσαρμογή των εργαλείων επεξεργασίας της Ελληνικής.

2005 – 2007. *BulTM: Μεταφραστική Πλατφόρμα από τα Ελληνικά στα Βουλγαρικά και αντίστροφα (INTERREG- PHARE CBC – ΕΛΛΑΔΑ – ΒΟΥΛΓΑΡΙΑ)*. Στόχος του έργου ήταν η ανάπτυξη μιας πλατφόρμας που θα διευκολύνει τη μεταφραστική διαδικασία από τα ελληνικά στα βουλγαρικά και αντίστροφα. Ο μεθοδολογικός του στόχος ήταν να χρησιμοποιηθούν σύγχρονες μέθοδοι επεξεργασίας παράλληλων κειμένων με στόχο την ανάπτυξη μεταφραστικών μνημών. Το σύστημα που δημιουργήθηκε συνδυάζει διαφορετικά επίπεδα επεξεργασίας και πληροφορίας, είναι ιδιαίτερα αποδοτικό ενώ μπορεί να χρησιμοποιείται στη μεταφραστική διαδικασία στα βουλγαρικά και ελληνικά σε θεματικές περιοχές που ενδιαφέρουν ιδιαίτερω τους πληθυσμούς των διασυνοριακών περιοχών. Συμμετείχα σε όλα τα στάδια του έργου, από τον αρχικό σχεδιασμό και την υποβολή της πρότασης μέχρι την τελική υλοποίηση.

2004 – 2007. *REVEAL THIS: Retrieval of Video & Language for the Home user in an Information Society. (IST)*. Το έργο REVEAL THIS στόχευε στην έρευνα και ανάπτυξη ενός συστήματος *γλωσσικής επεξεργασίας* πολυμεσικού και πολυγλωσσικού περιεχομένου, δηλαδή εικόνων βίντεο αλλά και κειμένου σε διάφορες γλώσσες, με στόχο την κατά το δυνατόν ακριβέστερη δεικτοδότηση, ευρετηρίαση και ανάκτησή τους. Η έρευνα επικεντρώθηκε σε δύο γλώσσες (Αγγλικά και Ελληνικά), και στις θεματικές κατηγορίες *ΠΟΛΙΤΙΚΗ* (συνεδριάσεις του Ευρωπαϊκού Κοινοβουλίου, συναφείς συνεντεύξεις τύπου και προγράμματα ειδήσεων ευρωπαϊκών τηλεοπτικών σταθμών) και *τουρισμός* (ταξιδιωτικά προγράμματα, κυρίως ντοκιμαντέρ και ενημερωτικές εκπομπές για διάφορους γεωγραφικούς τόπους αλλά και κείμενα τουριστικού περιεχομένου). Το τελικό σύστημα που δημιουργήθηκε (prototype) επιτρέπει διαγλωσσική αναζήτηση περιεχομένου μέσω ιστοτόπου, καθώς και αυτόματη προώθηση επιλεγμένου περιεχομένου στο κινητό τηλέφωνο του χρήστη με βάση τα προσωπικά του ενδιαφέροντα. Το REVEAL THIS συνδυάζει τεχνολογίες επεξεργασίας *φωνής, κειμένου, εικόνας και video* (αυτόματη μεταγραφή προφορικού λόγου σε γραπτό, αναγνώριση Ονοματικών Οντοτήτων, εξαγωγή όρων και γεγονότων, αναγνώριση προσώπων και ομιλητών, κατηγοριοποίηση εικόνων και κειμένου, αυτόματη μετάφραση) για τις λειτουργίες δεικτοδότησης/αναζήτησης, φιλτραρισμένης προώθησης και περίληψης πολυμεσικού περιεχομένου.

Στο πλαίσιο του έργου συμμετείχα στις εργασίες για την επεξεργασία της Ελληνικής, και την ανάπτυξη και προσαρμογή των σχετικών υπολογιστικών εργαλείων (εργαλείο μορφοσυντακτικής ανάλυσης και λημματοποίησης, συντακτικός αναλυτής, εργαλείο εντοπισμού ονοματικών οντοτήτων και χρονικών εκφράσεων), και την ανάπτυξη των ηλεκτρονικών σωμάτων κειμένων για την αξιολόγηση των παραπάνω συστημάτων. Επίσης, στα πλαίσια του έργου συμμετείχα στις εργασίες για την ανάπτυξη ηλεκτρονικού συντακτικού λεξικού για την Ελληνική στο πρότυπο του PropBank.

2005 – 2006. “Συγγραφή Νέων Βιβλίων και Παραγωγή Υποστηρικτικού Εκπαιδευτικού Υλικού με βάση το ΔΕΠΠΣ και τα ΑΠΣ για το Γυμνάσιο”. ΕΠΕΑΕΚ II 2004-2006. Συμμετοχή στην συγγραφική ομάδα του Ερμηνευτικού Λεξικού για το Γυμνάσιο που εκπονήθηκε στο πλαίσιο του Προγράμματος του Παιδαγωγικού Ινστιτούτου σε συνεργασία με τον Εκδοτικό Οίκο ΜΕΤΑΙΧΜΙΟ.

2005 – 2006. ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ – ΕΠΑΝ: EXCELAN: Ενίσχυση της αριστείας ερευνητικών και τεχνολογικών φορέων εποπτευόμενων από την Γ.Γ.Ε.Τ. Αριστεία στην Γλωσσική Τεχνολογία. Αναπληρώτρια επιστημονική υπεύθυνη. Στο πλαίσιο του έργου, εργάστηκα για τη βελτίωση, επέκταση και προσαρμογή του εργαλείου Αναγνώρισης Ονοματικών Οντοτήτων του ΙΕΛ, με στόχο την κάλυψη περισσότερων γλωσσικών φαινομένων σύμφωνα με τις πρόσφατες προδιαγραφές (Automatic Content Extaction).

2003 – 2004. eContent-Intera: Integrated European Resource Area. Το έργο INTERA είχε δύο άξονες: (i) τη διαμόρφωση ενιαίου χώρου Ευρωπαϊκών Γλωσσικών Πόρων (με τη σύνδεση διεθνών, εθνικών και τοπικών κέντρων που ασχολούνται με γλωσσικά δεδομένα, και (ii) την παραγωγή νέων πολυγλωσσικών πόρων. Ο πρώτος στόχος αφορούσε την ομογενοποίηση διαφορετικών ειδών γλωσσικών πόρων με τη βοήθεια μετα-δεδομένων, και τη διασύνδεση των πόρων με τον ανάλογο κατάλογο υπολογιστικών εργαλείων επεξεργασίας τους, έτσι ώστε να είναι δυνατή η χρήση των κατάλληλων εργαλείων για τους πόρους. Το INTERA στόχευε, με τη δημιουργία αυτού του ενιαίου χώρου, να διευκολύνει την πρόσβαση στους Ευρωπαϊκούς γλωσσικούς πόρους, να βοηθήσει τη βιομηχανία, την έρευνα και την εκπαίδευση σχετικά με το ηλεκτρονικό περιεχόμενο και να αυξήσει τη χρήση των υπάρχοντων πόρων. Με τον δεύτερο στόχο, το έργο είχε σκοπό να αντιμετωπίσει την έλλειψη ποιοτικών πολυγλωσσικών πόρων, κυρίως για τις λιγότερο ομιλούμενες γλώσσες, συμπεριλαμβανομένων και των Βαλκανικών, οι οποίες είναι μεγάλης σημασίας για την βιομηχανία του ηλεκτρονικού περιεχομένου. Στο πλαίσιο του έργου συμμετείχα (α) στην ανάπτυξη γλωσσικών πόρων (σωμάτων κειμένων – λεξικών ορολογίας), και (β) στη μελέτη και εναρμόνιση των συνόλων χαρακτηριστικών (tagsets) που είναι απαραίτητα για τη μορφοσυντακτική ανάλυση των γλωσσών του έργου. Αποτέλεσμα της εργασίας αυτής υπήρξε η εναρμόνιση υφιστάμενων tagsets (εν προκειμένω, των EAGLES και MULTEXT-East – καθένα εκ των οποίων είχε δημιουργηθεί για να καλύψει διαφορετικές γλώσσες), και ο καθορισμός ενός κοινά αποδεκτού tagset (INTERA tagset), το οποίο καλύπτει τυπολογικά διαφορετικές γλώσσες. Σήμερα, οι πόροι που δημιουργήθηκαν στο έργο INTERA (επισημειωμένα και στοιχισμένα ηλεκτρονικά σώματα κειμένων, υπολογιστικά λεξικά, και το εναρμονισμένο INTERA tagset) αποτελούν γλωσσικούς πόρους αναφοράς για νεότερα εργαλεία επεξεργασίας των εμπλεκόμενων γλωσσών – μεταξύ των οποίων και της Ελληνικής.

IST-Cimwos: Combined Image and Word Spotting. Το έργο αποσκοπούσε στη διευκόλυνση της αρχειοθέτησης και ανάκτησης οπτικοακουστικού υλικού. Στόχος ήταν η ανάπτυξη και ενσωμάτωση ενός εύρωστου αλγόριθμου εντοπισμού λέξεων-κλειδιών και ενός αποδοτικού αλγόριθμου εντοπισμού εικόνων ειδικά σχεδιασμένων για ψηφιακό οπτικοακουστικό περιεχόμενο, που θα οδηγήσει στην υλοποίηση και επίδειξη ενός συστήματος (prototype) για αποδοτική ανάκτηση από πολυμεσικές βάσεις δεδομένων. Ειδικότερα, αναπτύχθηκε ένα σύστημα για την αυτόματη ανάκτηση εικόνων, βίντεο και ομιλίας από μια οπτικοακουστική βάση δεδομένων σύμφωνα με λέξεις-κλειδιά που εισάγει ο χρήστης. Ο συνδυασμός εντοπισμού λέξεων και εικόνων που χρησιμοποιήθηκε επιτρέπει εστιασμένες και ακριβείς αναζητήσεις με

αυξημένη λειτουργικότητα και ανθεκτικότητα. Συμμετοχή στην ανάπτυξη γλωσσικών πόρων (επισημειωμένων σωμάτων κειμένων) και σχετικών προδιαγραφών.

IST – HLT: Balric-Ling: Balkan Regional Information Centers for Awareness and Standardization of Linguistic Resources and Tools for Advanced HLT Applications. Στόχος του BALRIC-LING ήταν η ενημέρωση και κινητοποίηση των υποψηφίων, τότε, μελών της Ευρωπαϊκής Ένωσης από τα Βαλκάνια (Βουλγαρία και Ρουμανία) σχετικά με τις δυνατότητες των προηγμένων γλωσσικών τεχνολογιών και τη χρήση τους σε επιστημονικές και βιομηχανικές εφαρμογές. Αποτέλεσε έργο μεγάλης σημασίας για τις χώρες αυτές, καθώς μολονότι σύντομα θα εισχωρούσαν επισήμως στην Ευρωπαϊκή Κοινότητα, δεν είχαν να επιδείξουν παραδείγματα προώθησης εφαρμογών γλωσσικής τεχνολογίας. Έχοντας μείνει απομονωμένες για χρόνια από το επιστημονικό γίγνεσθαι της Δυτικής Ευρώπης, και έχοντας να αντιμετωπίσουν διαφορετικές γλώσσες, οι ελάχιστες ομάδες γλωσσικής τεχνολογίας και στις δύο χώρες αδυνατούσαν ακόμα να συγκεντρώσουν από μόνες τους ειδικούς που θα εργάζονταν με στόχο να βελτιώσουν την ποιότητα των σχετικών εφαρμογών στις μόλις αναπτυσσόμενες αγορές γλωσσικής τεχνολογίας στη Βουλγαρία και τη Ρουμανία. Το έργο εστίασε στα εξής: (α) γλωσσικοί πόροι (λεξικά και σώματα κειμένων), (β) γλωσσικός σχολιασμός, (γ) εργαλεία γλωσσικού σχολιασμού, και (δ) πιθανές εφαρμογές των (α) και (β). Αναπληρώτρια επιστημονική υπεύθυνη στο ΙΕΛ.

ΕΠΕΤ-II-ΚΑΙΡΟΣ: Αυτόματη Εξαγωγή Προγνωστικών Δελτίων Καιρού. Πρόγραμμα Ενίσχυσης Ερευνητικού Δυναμικού (ΠΕΝΕΔ99). Το έργο ΚΑΙΡΟΣ απέβλεπε στην αυτόματη σύνθεση και παραγωγή σχεδίων δελτίων καιρού (εφεξής ΔΚ), από αριθμητικά μετεωρολογικά δεδομένα, σε δύο γλώσσες, Αγγλικά και Ελληνικά. Τα αποτελέσματα του έργου αυτού θα δώσουν στους ερευνητές στον τομέα της γλωσσικής τεχνολογίας τη δυνατότητα να χρησιμοποιήσουν την απαραίτητη υποδομή και τεχνογνωσία για παραγωγή φυσικής γλώσσας (natural language generation, εφεξής ΠΦΓ) που δημιουργήθηκε κατά την εκτέλεση του έργου. Τελικό προϊόν του έργου ΚΑΙΡΟΣ αποτελούν ένα πρότυπο σύστημα αυτόματης παραγωγής σχεδίων ΔΚ, ένα σύνολο γλωσσικών πόρων, προσανατολισμένων στην υπογλώσσα (sublanguage) της μετεωρολογίας και κλιματολογίας για δύο φυσικές γλώσσες (Αγγλικά και Ελληνικά), καθώς και μια μεθοδολογία σύνθεσης διαφόρων τύπων δελτίων με τη χρήση προτύπων XML, βασισμένη σε σύγχρονες αρχιτεκτονικές. Συμμετοχή στην ανάπτυξη γραμματικών.

ΕΠΕΤ-II-Οικονομία: Επιφανειακή Κατανόηση Κειμένων για Αποδοτική Δεικτοδότηση και Εξαγωγή Πληροφοριών από Οικονομικά Δεδομένα. Πρόγραμμα Ενίσχυσης Ερευνητικού Δυναμικού (ΠΕΝΕΔ99). Το έργο Οικονομία απέβλεπε στην ανάπτυξη εύρωστων μεθόδων επιφανειακής κατανόησης ελεύθερου κειμένου. Συγκεκριμένα, αποσκοπούσε στην κατασκευή ενός συστήματος αυτόματης ανάλυσης για την παραγωγή από ελεύθερο κείμενο μιας επιφανειακής σημασιολογικής αναπαράστασης, που περιλαμβάνει πληροφορία γλωσσική (μορφοσυντακτική ανάλυση και λήμμα για κάθε λέξη, αναγνώριση ονοματικών οντοτήτων και κατηγοριοποίηση αυτών, ανάλυση της επιφανειακής συντακτικής δομής κάθε πρότασης,

επίλυση συναναφοράς μεταξύ ονομάτων, ονοματικών φράσεων και αντωνυμικών τύπων). Η αναπαράσταση αυτή μπορεί να χρησιμοποιηθεί για την αποδοτική δεικτοδότηση κειμένων η οποία βασίζεται όχι μόνο στην "επίπεδη" πληροφορία που παρέχουν οι συχνότητες των λέξεων - κλειδιών, αλλά και στην σύνθετη επεξεργασία - σε γραμματικό και συντακτικό επίπεδο - της αναζήτησης του χρήστη. Αναπλήρωτρια επιστημονική υπεύθυνη του έργου.

SOCRATES/LINGUA (TM-LD-1995-1-GR-30): Electronic Dictionary of Pronunciation and Usage of Modern Greek for Foreigners. ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ - ΕΡΓΑΣΤΗΡΙΟ ΕΝΣΥΡΜΑΤΗΣ ΤΗΛΕΠΙΚΟΙΝΩΝΙΑΣ (ΕΕΤ), (Συντονιστής), BRASSHOUSE LANGUAGE CENTRE, BIRMINGHAM (UK). Στόχος του έργου ήταν η δημιουργία εύχρηστου ηλεκτρονικού λεξικού (45.000 λήμματα + 10.000 παραδείγματα χρήσης) για ξένους μαθητές και χρήστες της ελληνικής με καινοτόμα χαρακτηριστικά. Η ανάπτυξη του λεξικού πραγματοποιήθηκε με βάση το Σώμα Κειμένων του Εργαστηρίου Ενσύρματης Τηλεπικοινωνίας (Corpus based lexicography), ενώ η πιλοτική χρησιμοποίηση του λεξικού στο BRASSHOUSE είχε ως αποτέλεσμα τον εκτεταμένο έλεγχο και διόρθωση του λεξικού. Στα χαρακτηριστικά του συγκαταλέγονται τα ακόλουθα: (α) φωνητική απόδοση των λημμάτων και μέρους των παραδειγμάτων (ηχογραφήσεις), (β) φωνητική γραφή κάθε λήμματος (διεθνές φων. αλφάβητο), (γ) γραμματική κατηγορία με παραπομπή σε κλιτικό παράδειγμα, (δ) πρόσβαση σε λήμμα από μορφή λήμματος, (ε) πρόσβαση με ορθή και εσφαλμένη ορθογραφική γραφή λήμματος, (στ) πρόσβαση με φωνητική γραφή (παροχή εναλλακτικών ορθογραφικών μορφών από φωνητικό λεξικό), (ζ) αντίστροφο λεξικό (ταξινόμηση λημμάτων με βάση την κατάληξη), (η) πίνακες τοπωνυμίων και ονομάτων.

HARPER-COLLINS: Bilingual Greek-English and English-Greek Electronic Dictionary (1993-1994). HarperCollins Publishers, Ltd, UK (1993-1994). Στόχος του έργου ήταν η δημιουργία δίγλωσσου ελληνο-αγγλικού και αγγλο-ελληνικού λεξικού με τη χρήση Σωμάτων Κειμένων (Corpus-based Lexicography). Στο πλαίσιο του έργου αναπτύχθηκε ισορροπημένο (balanced) Σώμα κειμένων αναφοράς (reference corpus) της ελληνικής γλώσσας του Εργαστηρίου Ενσύρματης Τηλεπικοινωνίας (60 εκ. λέξεων). Επίσης, αναπτύχθηκε το λογισμικό υπολογιστικής λεξικογραφίας "Λεξιθήρας", το οποίο επιτρέπει την επεξεργασία πολύ μεγάλων σωμάτων κειμένων φυσικής γλώσσας, η οποία συνίσταται στην ευρεία στατιστική ανάλυση των κειμένων, ληματοποίηση, γραμματική ανάλυση λέξεων κ.α. και παρέχει στο λεξικογράφο την δυνατότητα να επιλέξει κατάλληλα παραδείγματα χρήσεως.

ONOMASTICA (LRE-61004). Multi-Language Pronunciation Dictionary of Proper Names and Place Names. Funded by the European Union (1993-1994). Εργάστηκε για την κωδικοποίηση της προφοράς των κυρίων ονομάτων της Ελληνικής με τη χρήση του Διεθνούς Φωνητικού Αλφαβήτου.

NOW - Λεξικογραφία με Χρήση Υπολογιστών (1993-1995). Εκπαίδευση στην ηλεκτρονική λεξικογραφία και συμμετοχή στη λεξικογραφική εργασία του Εργαστηρίου (δίγλωσσο λεξικό της Νέας Ελληνικής, λεξικό ιδιωμάτων της Κάτω Ιταλίας).

ΥΠΟΒΟΛΗ ΠΡΟΤΑΣΕΩΝ ΓΙΑ ΧΡΗΜΑΤΟΔΟΤΗΣΗ ΕΡΓΩΝ Ε&Α (εγκεκριμένα έργα)

Υπολογιστικές Επιστήμες και Τεχνολογίες Δεδομένων, Περιεχομένου και Αλληλεπίδρασης, Δράση «Στρατηγική Ανάπτυξης Ερευνητικών και Τεχνολογικών Φορέων» του Επιχειρησιακού Προγράμματος Επιχειρηματικότητα Ανταγωνιστικότητα και Καινοτομία (ΕΠΑΝΕΚ).

ΠΟΛΥΤΡΟΠΟΝ (Αναπτυξιακές Προτάσεις Ερευνητικών Φορέων – ΚΡΗΠΙΣ, Υπουργείο Παιδείας, Διά Βίου Μάθησης & Θρησκευμάτων – Γ.Γ.Ε.Τ.).

FP7 – ICT. Analysis and Evaluation of Comparable Corpora for Under-Resourced Areas of Machine Translation (ACCURAT).

INTERREG- PHARE CBC – ΕΛΛΑΔΑ – ΒΟΥΛΓΑΡΙΑ. BuITM: Μεταφραστική Πλατφόρμα από τα Ελληνικά στα Βουλγαρικά και αντίστροφα.

TV++. Ε.Π “Κοινωνία της Πληροφορίας”, Μέτρο 3.3 “Έρευνα και Τεχνολογική Ανάπτυξη στην Κοινωνία της Πληροφορίας”, Πρ. “ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΩΝ, ΗΧΟΥ ΚΑΙ ΓΛΩΣΣΑΣ”.

MUSE: Multimedia Content Management System. Γενική Γραμματεία Έρευνας – Τεχνολογίας, e-business.

‘Πανόπτης’ Σύστημα Ψηφιακής Εποπτείας P/T Μ.Μ.Ε. Ε.Π “Κοινωνία της Πληροφορίας”, Μέτρο 2.2. “Ηλεκτρονική κυβέρνηση για την εξυπηρέτηση του πολίτη”.

eClassify. Σύστημα Διαχείρισης Περιεχομένου με υποστήριξη αυτόματης κατηγοριοποίησης αρχείων. Γ’ Κοινωνικό Πλαίσιο Στήριξης, Ε.Π. ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ, Μέτρο 4.3.

ΕΠΙΒΛΕΨΗ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ

Ηφαιστίων Δημήτριος Χριστόπουλος. 2023. Ανάλυση των δοξαστικών ρημάτων και των ρημάτων έκφρασης της Ελληνικής στη βάση της σημασιολογίας πλαισίων. Δημιουργία λεξικού πόρου και επισημείωση σωμάτων κειμένων. Μεταπτυχιακή Εργασία. ΠΜΣ Μετάφρασης και Διερμηνείας. Τμήμα Τουρκικών και Σύγχρονων Ασιατικών Σπουδών. Σχολή Οικονομικών και Πολιτικών Επιστημών. Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.

Γιαννοπούλου Στέλλα. 2019. Μεταναστευτική και προσφυγική κρίση. Σημασιολογικός σχολιασμός του θεματικού πεδίου μέσω των μεθοδολογιών της Γλωσσολογίας Σωμάτων Κειμένων. Μεταπτυχιακή Εργασία. ΠΜΣ Μετάφρασης και Διερμηνείας. Τμήμα Τουρκικών και Σύγχρονων Ασιατικών Σπουδών. Σχολή Οικονομικών και Πολιτικών Επιστημών. Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.

Παπαδέλλη Στέλλα. 2018. Αναγνώριση ρηματικών πολυλεκτικών σε Σώματα Κειμένων για την υποβοήθηση της μεταφραστικής διαδικασίας. Μεταπτυχιακή Εργασία. ΠΜΣ Μετάφρασης και Διερμηνείας. Τμήμα Τουρκικών και Σύγχρονων Ασιατικών Σπουδών. Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.

Πιλιτσιδίου Βέρα. 2018. Η ορολογία της Τραπεζικής. Ορογραφία στο πλαίσιο της Σημασιολογίας Πλαισίων. Μεταπτυχιακή Εργασία. ΠΜΣ Μετάφρασης και Διερμηνείας. Τμήμα Τουρκικών και Σύγχρονων Ασιατικών Σπουδών. Σχολή Οικονομικών και Πολιτικών Επιστημών. Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.

Κρεμμύδα Ναταλία. 2018. Δημιουργία δίγλωσσου ορολογικού πόρου στο θεματικό πεδίο των εμπορικών συναλλαγών για την υποβοήθηση της μεταφραστικής διαδικασίας στο ζεύγος γλωσσών Ελληνικής – Τουρκικής με εφαρμογή της μεθοδολογίας σωμάτων κειμένων σε συγκρίσιμο υλικό. Μεταπτυχιακή Εργασία. ΠΜΣ Μετάφρασης και Διερμηνείας. Τμήμα Τουρκικών και Σύγχρονων Ασιατικών Σπουδών. Σχολή Οικονομικών και Πολιτικών Επιστημών. Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.

Louizou, Sevasti. 2013. Development of a Computational Bilingual (Greek – Swedish) Lexicon of Emotion words from existing resources. Semantic and syntactic considerations. Masters in Language technology, University of Gothenburg.

ΕΠΙΚΟΥΡΙΑ ΕΡΕΥΝΗΤΩΝ ΙΕΛ ΓΙΑ ΤΗΝ ΕΠΙΒΛΕΨΗ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ

Καζανά, Ευαγγελία και Παπαμιχαλάκη Μαρία. 2013. Λεκτικά Ρήματα και επισημείωση γεγονοτικότητας. Διπλωματική Εργασία. ΤΕΧΝΟΓΛΩΣΣΙΑ: Διεπιστημονικό - Διαπανεπιστημιακό Μεταπτυχιακό Πρόγραμμα Ειδίκευσης στην Γλωσσική Τεχνολογία. (Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Εθνικό Μετσόβιο Πολυτεχνείο).

Μαρζέλου, Ευριδίκη και Μαρία Ζουράρη. 2008. Κειμενική Συνεπαγωγή. Από την θεωρία στην πράξη. Διπλωματική Εργασία. ΤΕΧΝΟΓΛΩΣΣΙΑ: Διεπιστημονικό - Διαπανεπιστημιακό Μεταπτυχιακό Πρόγραμμα Ειδίκευσης στην Γλωσσική Τεχνολογία. (Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Εθνικό Μετσόβιο Πολυτεχνείο).

Δεσύπρη, Ευαγγελία και Ειρήνη Σπυροπούλου. 2000. Συναναφορά στη Νέα Ελληνική. Διπλωματική Εργασία. ΤΕΧΝΟΓΛΩΣΣΙΑ: Διεπιστημονικό - Διαπανεπιστημιακό Μεταπτυχιακό Πρόγραμμα Ειδίκευσης στην Γλωσσική Τεχνολογία. (Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Εθνικό Μετσόβιο Πολυτεχνείο).

ΕΠΙΒΛΕΨΗ ΠΡΑΚΤΙΚΗΣ ΑΣΚΗΣΗΣ ΦΟΙΤΗΤΩΝ

Σεντούμπε – Τσούτση, Αθηνά. 2022. Μελέτη και επισημείωση προσβλητικού λόγου και ρητορικής μίσους σε κοινωνικά δίκτυα. (Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, τμήμα Γερμανικής Γλώσσας και Φιλολογίας).

VI. ΛΟΙΠΗ ΕΡΕΥΝΗΤΙΚΗ ΔΡΑΣΤΗΡΙΟΤΗΤΑ

ΜΕΛΟΣ ΕΠΙΣΤΗΜΟΝΙΚΗΣ ΕΠΙΤΡΟΠΗΣ ΣΥΝΕΔΡΙΩΝ ΚΑΙ ΘΕΜΑΤΙΚΩΝ ΗΜΕΡΙΔΩΝ

EMNLP 2024 System Demonstration Track (EMNLP24 Demo).

The 2024 Conference on Empirical Methods in Natural Language Processing (Industry Track). EMNLP 2024 Industry Track.

Sixth edition of the biannual International Conference on Computational Linguistics in Bulgaria (CLIB 2024), 9 – 10 September 2024 in Sofia, Bulgaria

21st EURALEX International Congress Lexicography and Semantics, 8-12 October 2024, Cavtat/Dubrovnik, Croatia.

2nd Workshop on Resources and Technologies for Indigenous, Endangered and Lesser-resourced Languages in Eurasia. Turin, Italy, May 25, 2024. Collocated with LREC-COLING 2024.

The 9th Workshop on Linked Data in Linguistics: Resources, Applications, Best Practices (LDL 2024). Turin, Italy, May 25, 2024, Torino, Italy and online. Collocated with LREC-COLING 2024.

2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Mexico City, Mexico. June 16–21, 2024

*SEM 2024: The 13th Joint Conference on Lexical and Computational Semantics. Co-located with ACL 2023 (Mexico City, Mexico), June 16.

The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation LREC-COLING 2024, Turin, Italy, 20-25 May 2024.

The 2023 Conference on Empirical Methods in Natural Language Processing - Industry Track - EMNLP Industry Track 2023, Singapore, December 6 –10, 2023.

The 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023), Singapore, December 6 –10, 2023.

The 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (IJCNLP-AAACL 2023), Bali, Indonesia, November 1-4, 2023.

*SEM 2023: The 12th Joint Conference on Lexical and Computational Semantics, Co-located with ACL 2023 (Toronto, Canada), July 13-14, 2023.

61st Annual Meeting of the Association for Computational Linguistics 2023 (ACL'23), July 9 – 14, 2023, Toronto, Canada.

8th biennial conference on Electronic Lexicography in the 21st century (ELEX 2023): Invisible lexicography: everywhere lexical data is used without users realizing they make use of a “dictionary”. Brno, Czech Republic, 27–29 June 2023.

17th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2023), May 2 – 6, 2023, Dubrovnik, Croatia.

Association for Computational Linguistics ([ACL Rolling Review \(ARR\) 2021-2022](#))

*SEM 2022: The 11th Joint Conference on Lexical and Computational Semantics *SEM 2022, Seattle, Washington, 14-15 July 2022.

18th Workshop on Multiword Expressions (MWE 2022). Organized as LREC workshop, Marseille, France.

5th International Conference on Computational Linguistics in Bulgaria 2022 (CLIB 2022). 8 – 9 September 2022, Sofia, Bulgaria.

60th Annual Meeting of the Association for Computational Linguistics, 2022 (ACL'22), 22 – 27 May 2022, Dublin, Ireland.

13th International Conference on Language Resources and Evaluation, LREC 2022, 13-15 May 2020, Marseille, France.

7th biannual Conference on Electronic lexicography in the 21st century (ELEX 2021): post-editing lexicography. Brno, Czech Republic, July 5-7, 2021.

17th Workshop on Multiword Expressions (MWE 2021). Collocated with ACL-IJCNLP 2021 (Bangkok, Thailand), 5 or 6 August 2021

59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021).

*SEM 2021: The Tenth Joint Conference on Lexical and Computational Semantics *SEM 2021, August 5-6, 2021.

*SEM 2020: The Nineth Joint Conference on Lexical and Computational Semantics *SEM 2020, 12-13 December 2020, Barcelona, Spain.

Joint Workshop on Multiword Expressions and Electronic Lexicons (MWE-LEX 2020), co-located with COLING 2020, 13 December 2020, Barcelona, Spain.

58th Annual Meeting of the Association for Computational Linguistics (ACL'20). July 5-10, 2020. Seattle, Washington, USA.

12th International Conference on Language Resources and Evaluation, LREC 2020, 13-15 May 2020, Marseille, France.

Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019), 2 August 2019, Florence, Italy.

Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018), August 25-26, 2018. Workshop at COLING 2018, Santa Fe, USA.

11th International Conference on Language Resources and Evaluation, LREC 2018, 7-12 May 2018, Miyazaki (Japan).

Computational and Corpus-based Phraseology. Recent advances and interdisciplinary approaches. EUROPHRAS 2017 International Conference. 13-14 November 2017.

13th Workshop on Multiword Expressions (MWE 2017), Workshop at EACL 2017, April 4, 2017, Valencia, Spain.

10th International Conference on Language Resources and Evaluation, LREC 2016. 23-28 May, Portorož Slovenia.

9th International Conference on Language Resources and Evaluation, LREC 2014. 26-31 May, Reykjavik, Iceland.

Workshop on Annotation of Corpora for Research in the Humanities (ACRH-3), 12 December 2013, University of Sofia, Bulgaria. Co-located with the 12th International Workshop on Treebanks and Linguistic Theories" (TLT-12).

Το σημασιολογικό πεδίο των συναισθημάτων: δια-θεματικές και δια-γλωσσικές προσεγγίσεις. Από την λέξη στη φράση και το κείμενο. Θεματική συνεδρία στο πλαίσιο του 11^{ου} Διεθνούς Συνεδρίου Ελληνικής Γλωσσολογίας, Ρόδος.

International Symposium on Open-Source Intelligence & Web Mining 2012 (OSINT-WM 2012) in Conjunction with (European ISI 2012). 22-23 August 2012, Odense, Denmark.

8th *International Conference on Language Resources and Evaluation*, LREC 2012. 21 - 27 May 2012, Istanbul, Turkey.

Second *Workshop on Annotation and Exploitation of Parallel Corpora* (AEP2). RANLP 2011 Workshop, 15th or 16th September 2011, Hissar, Bulgaria

Computational Linguistics – Applications (CLA'11) Workshop, International Multi-conference on Computer Science and Information Technology, 17-19 October 2011, Warsaw, Poland.

Computational Linguistics – Applications (CLA'10) Workshop, International Multi-conference on Computer Science and Information Technology, 18-20 October 2010, Wisła, Poland.

48th *Annual Meeting of the Association for Computational Linguistics* (ACL 2010), Uppsala, Sweden, July 11–16, 2010

7th *International Conference on Language Resources and Evaluation*, LREC 2010. 17 - 23 May 2010, Valletta, Malta.

6th *International Conference on Language Resources and Evaluation*, LREC 2008. 26 May – 1 June 2008, Marrakech, Morocco.

ΜΕΛΟΣ ΟΡΓΑΝΩΤΙΚΗΣ ΕΠΙΤΡΟΠΗΣ ΣΥΝΕΔΡΙΩΝ ΚΑΙ ΕΠΙΣΤΗΜΟΝΙΚΩΝ ΔΙΑΓΩΝΙΣΜΩΝ

LREC-COLING 2024. Area Chair: LC03 - Corpora and Annotation (including Tools, Systems, Treebanks), and LC10 - Information Extraction, Knowledge Extraction, and Text Mining. Turin, Italy.

Joint Workshop on Multiword Expressions and Universal Dependencies (MWE-UD 2024). Collocated with LREC-COLING 2024 (Torino, Italia), May 25, 2024. Organised and sponsored by The Special Interest Group on the Lexicon (SIGLEX) of the Association for Computational Linguistics (ACL), SIGLEX's Multiword Expressions Section (SIGLEX-MWE), Universal Dependencies (UD) and UniDive Cost Action CA2116.

19th edition of the Workshop on Multiword Expressions (MWE 2023) Organized and sponsored by SIGLEX, the Special Interest Group on the Lexicon of the ACL. Full-day workshop collocated with EACL 2023, Dubrovnik, Croatia, May 2 or 6, 2023. (General Chair).

Shared task on semi-supervised identification of verbal multiword expressions - edition 1.2. Organized as part of the MWE-LEX 2020 workshop co-located with COLING 2020 (Barcelona, Spain), December 13, 2020.

Shared task on automatic identification of verbal multiword expressions - edition 1.1. Organized as part of the LAW-MWE-CxG 2018 workshop co-located with COLING 2018 (Santa Fe, USA), August 25-26, 2018 (Language Leader and Language Group Leader)

Shared task on automatic identification of verbal multiword expressions. Organized as part of the MWE 2017 workshop co-located with EACL 2017 (Valencia, Spain), April 3 or 4, 2017.

Το σημασιολογικό πεδίο των συναισθημάτων: δια-θεματικές και δια-γλωσσικές προσεγγίσεις. Από την λέξη στη φράση και το κείμενο. Θεματική συνεδρία στο πλαίσιο του 11^{ου} Διεθνούς Συνεδρίου Ελληνικής Γλωσσολογίας, Ρόδος.

7^ο Συνέδριο Ελληνική Γλώσσα και Ορολογία, Αθήνα 22-24 Οκτωβρίου 2009.

ΚΡΙΣΕΙΣ ΣΕ ΠΕΡΙΟΔΙΚΑ - EDITORIAL BOARD MEMBER

Frontiers in Artificial Intelligence. Language and Computation. [Review Editor](#).

International Journal of Lexicography. Oxford University Press (Reviewer)

Dictionaries: Journal of the Dictionary Society of North America (Reviewer)

Linguistics and Literature Studies Journal. (Editorial Board, 2012 - 2017).

ΘΕΣΕΙΣ ΕΥΘΥΝΗΣ ΣΕ ΕΠΑΓΓΕΛΜΑΤΙΚΕΣ ΚΑΙ ΕΠΙΣΤΗΜΟΝΙΚΕΣ ΕΝΩΣΕΙΣ

2022 – 2024: Εκλεγμένο μέλος του SIGLEX (ACL Special Interest Group on the Lexicon) ως εκπρόσωπος του MWE Section (Multiword Expressions).

2022 – 2024: Επικεφαλής του SIGLEX – MWE Section.

ΕΚΔΟΤΙΚΟ ΕΡΓΟ – ΕΠΙΣΤΗΜΟΝΙΚΕΣ ΔΗΜΟΣΙΕΥΣΕΙΣ ΑΝΟΙΚΤΗΣ ΠΡΟΣΒΑΣΗΣ

[Language Science Press](#). Editor-in-chief of the newly accepted book series *Frame-based Approaches to Semantics*.

ΣΥΜΜΕΤΟΧΗ ΣΕ ΕΠΑΓΓΕΛΜΑΤΙΚΕΣ ΚΑΙ ΕΠΙΣΤΗΜΟΝΙΚΕΣ ΕΝΩΣΕΙΣ

European Association for Lexicography (EURALEX)

Association for Computational Linguistics (ACL)

Association for Computational Linguistics (ACL) Special Interest Group on Computational Semantics (SIGSEM)

Association for Computational Linguistics (ACL) Special Interest Group on the Lexicon (SIGLEX)

Association for Computational Linguistics (ACL) Multi-Word Expressions section of the Special Interest Group on the Lexicon (SIGLEX-MWE)

Association for Computational Linguistics (ACL) Special Interest Group on Language Technologies for the Socio-Economic Sciences and Humanities (SIGHUM)

European Language Resources Association (ELRA)

Societas Linguistica Europaea (SLE)

International Language and Law Association (ILLA)

ΟΜΙΛΙΕΣ

2023 Γιούλη, Β. Γλωσσικοί πόροι στην εποχή των μεγάλων δεδομένων: “It takes two to tango”. Ομιλία στο πλαίσιο των εσωτερικών παρουσιάσεων στο Ινστιτούτο Επεξεργασίας του Λόγου, ΕΚ ΑΘΗΝΑ (13.09.2023).

- 2022 Γιούλη, Β. Πολυλεκτικές Εκφράσεις και Υπολογιστική Επεξεργασία Κειμένου. Ομιλία στο πλαίσιο του μαθήματος *ΓΛΩΣΣΟΛΟΓΙΑ ΣΩΜΑΤΩΝ ΚΕΙΜΕΝΩΝ* του Προπτυχιακού Προγράμματος Σπουδών Πανεπιστημίου Αιγαίου (17.01.2022).
- 2014 Γιούλη, Β. Πίνακες του Λεξικού-Γραμματική και Τοπικές Γραμματικές για την ανάπτυξη εφαρμογών Επεξεργασίας Φυσικής Γλώσσας. Ομιλία στο πλαίσιο του μαθήματος *ΗΛΕΚΤΡΟΝΙΚΗ ΛΕΞΙΚΟΓΡΑΦΙΑ* του Μεταπτυχιακού Προγράμματος Ειδίκευσης στην Γλωσσική Τεχνολογία *ΤΕΧΝΟΓΛΩΣΣΙΑ* (04.2014).
- 2012 Γιούλη, Β. Παράλληλα και Συγκρίσιμα Σώματα Κειμένων: τυπολογία, συλλογή και χρήση στη Μηχανική Μετάφραση. Ομιλία στο πλαίσιο του μαθήματος *Μηχανική Μετάφραση* του Μεταπτυχιακού Προγράμματος Ειδίκευσης στην Γλωσσική Τεχνολογία *ΤΕΧΝΟΓΛΩΣΣΙΑ* (16.12.2012).
- 2011 Γιούλη, Β. Ανάλυση συναισθήματος. Λεξικογραφικές προσεγγίσεις. Ομιλία στο πλαίσιο του μαθήματος *Ηλεκτρονική Λεξικογραφία*, του Μεταπτυχιακού Προγράμματος Ειδίκευσης στην Γλωσσική Τεχνολογία *ΤΕΧΝΟΓΛΩΣΣΙΑ* (19.07.2011).
- 2011 Γιούλη, Β. Σώματα Συγκρίσιμων Κειμένων: Προσεγγίσεις και Πρακτικές Εφαρμογές στη Μηχανική Μετάφραση. Ομιλία στο πλαίσιο του μαθήματος *ΜΗΧΑΝΙΚΗ ΜΕΤΑΦΡΑΣΗ* του Μεταπτυχιακού Προγράμματος Ειδίκευσης στην Γλωσσική Τεχνολογία *ΤΕΧΝΟΓΛΩΣΣΙΑ* (21.06.2011).

VII. ΔΗΜΟΣΙΕΥΣΕΙΣ

Citations: 726, h-index 13, i10-index 15 (Google Μελετητής, Απρίλιος 2024)

BIBLIA

Φωτοπούλου, Α. και Γιούλη, Β. (Υπό έκδοση). Περί Λεξιλογίου. Λεξικο-λογία και Λεξικο-γραφία: έντυπη, ηλεκτρονική, υπολογιστική. Εκδόσεις ΚΑΛΛΙΠΟΣ.

ΣΥΛΛΟΓΙΚΟΙ ΤΟΜΟΙ (Επιμ.)

Giouli, V. and Barbu Mititelu, V. (Eds). (2024). *Multiword expressions in lexical resources. Linguistic, Lexicographic, and Computational perspectives*. Phraseology and Multiword Expressions (PMWE) series. Language Science Press.

Barbu Mititelu, V. and **Giouli, V.** (Eds). (In preparation). *Lexical resources in Natural Language Processing Tasks*. Phraseology and Multiword Expressions (PMWE) series. Language Science Press.

ΚΕΦΑΛΑΙΑ ΣΕ ΒΙΒΛΙΑ (με κριτές)

- [B01] Giouli, V., Pilitsidou, V. and Christopoulos, H. (2024). A FrameNet approach to deep semantics for MWEs. In Voula Giouli & Verginica Barbu Mititelu (eds.), *Multiword expressions in lexical resources: Linguistic, lexicographic, and computational perspectives*, pp. 147–186. Berlin: Language Science Press. DOI: 10.5281/zenodo.10998639
- [B02] Savary, A., Candito, M., Barbu Mititelu, V., Bejček, E., Cap, F., Čéplö, S., Cordeiro, S.R., Eryiğit, G., Giouli, V., van Gompel, M., Hacoheh-Kerner, Y., Kovalevskaite, J., Krek, S., Liebeskind, C., Monti, J., Parra Escartín, C., van der Plas, L., Qasemizadeh, B., Ramisch, C., Sangati, F., Stoyanova, I., and Vincze, V. 2018. PARSEME multilingual corpus of verbal multiword expressions. In Markantonatou, S., et al. (eds.), *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop*, pp. 87-147. Berlin: Language Science Press. DOI: 10.5281/zenodo.1471591.
- [B03] Fotopoulou, A. and Giouli, V. 2018. MWEs and the Emotion Lexicon: Typological and Cross-lingual considerations. In Sailer, M. and Markantonatou, S. (eds) *Multiword Expressions: Insights from a Multi-lingual Perspective*. pp. 63–91. Berlin: Language Science Press. DOI:10.5281/zenodo.1182591.
- [B04] Giouli, V., A. Fotopoulou, E. Mouka, and I.E. Saridakis. 2013. *Annotating Sentiment Expressions for Lexical Resources*. In Blumenthal, P., Novakova, I., Siepmann, D. (eds.), *Les émotions dans le discours. Emotions in discourse*, σσ. 281-295. Frankfurt, Main et al.: Peter Lang. DOI: 10.13140/RG.2.1.3827.3365
- [B05] Giouli, V., Simov, K., and Osenova, P. 2011. A parallel Greek-Bulgarian corpus: a digital resource of the shared Cultural Heritage. In Spolender, C., van de Bosch, A., and Zervanou, K. (eds.) *Language Technology for Cultural Heritage: Theory and Applications of Natural Language Processing Series*, pp. 99-114. Springer. DOI: 10.1007/978-3-642-20227-8_6
- [B06] Boutsis S., Demiros I., Giouli V., Liakata M., Papageorgiou H., Piperidis S. (2000) A System for Recognition of Named Entities in Greek. In: Christodoulakis D.N. (eds) *Natural Language Processing — NLP 2000. NLP 2000. Lecture Notes in Computer Science*, vol 1835. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45154-4_39.

ΠΕΡΙΟΔΙΚΑ

- [J01] Antoniou-Kritikou I., **Giouli V.**, Tsoulouhas G. and Economou C. (2024). Using LLMs in a Language Teaching/Learning Application. ERCIM News (136), January 2024.
- [J02] **Giouli, V.** 2023. A model for representing the semantics of MWEs: from lexical semantics to the semantic annotation of complex predicates. In *Frontiers in Artificial Intelligence, Language and Computation*, Volume 6 - 2023 | [doi: 10.3389/frai.2023.802218](https://doi.org/10.3389/frai.2023.802218)
- [J03] **Giouli, V.** 2020. The semantics of MWEs: representation in a computational lexical resource. In Markantonatou, S. and Christophidou, A. (Eds). *Multiword Expressions: Drawing on Data from Modern Greek and other languages*, *Bulletin of Scientific Terminology and Neologisms* (15) 193-218. Academy of Athens.
- [J04] **Giouli, V.**, and Markantonatou, S. 2019. State-of-the-art on monolingual lexicography for Greece (el). Στο *Slovenščina 2.0*, 7(1): 78–97. DOI: <http://dx.doi.org/10.4312/slo2.0.2019.1.78-97>.
- [J05] Savary, A., S. R. Cordeiro, T. Lichte, C. Ramisch, U. Iñurrieta, and V. **Giouli**. 2019. Literal Occurrences of Multiword Expressions: Rare Birds That Cause a Stir. *Prague Bulletin of Mathematical Linguistics* 112 (2019): 5-54. DOI: 10.2478/pralin-2019-0001

[J06] Kosem, I., Lew, R., Müller-Spitzer, C., Ribeiro Silveira, M., Wolfer, S., Dorn, A., Gurrutxaga, A., Ceberio, K., Etxeberria, A., Lefer, M.A., Geeraerts, D., Despot, K. Š., Stojanov, T., Ljubešić, N., Škrabal, M., Štěpánková, B., Vodrážková, V., Lorentzen, H., Trap-Jensen, L., Kallas, J., Tuulik, M., Koppel, K., Langemets, M., Heinonen, T., Thomas, I., Margilitadze, T., Markantonatou, S., **Giouli**, V., Mulhall, C., Kernerman, I., Ben-Moshe, Y., Sadan, T. Abel, A., Curcio, M.N., Tanturovska, L., Nikovska, B., Tiberius, C., Grønvik, O., Hovdenak, O.M., Berg-Olsen, S., Karlsen, K.E., Smith Ore, C.-E., Biesaga, M., Kuhn, T.Z., Silvestre, J., Tamba, E.I., Haja, G., Clim, M.-R., Patrascu, M.-I., Tasovac, T., Petrović, S., Holdt, Š.A., Rivei, V. 2018. The image of the monolingual dictionary across Europe. Results of the European Survey on Dictionary Use and Culture. *International Journal of Lexicography*, 32 (1), 92-114. DOI: 10.1093/ijl/ecy022.

[J07] **Giouli**, V. 2011. *Annotating corpora from various sources in the humanities domain. Shortcomings and issues*. *Journal for Language Technology and Computational Linguistics*, (JLCL), 26 (2). *Special issue on Annotating Corpora for Research in the Humanities*, pp. 53-63.

[J08] Skadina, I., Aker, A., **Giouli**, V., Tufis, D. Gaizauskas, R., Mierina, M., and Mastropavlos, N. 2010. A Collection of comparable corpora for under-resourced languages. In *Frontiers in Artificial Intelligence and Applications*, Vol. 219, pp. 161–168.

ΑΡΘΡΑ ΣΕ ΠΡΑΚΤΙΚΑ ΣΥΝΕΔΡΙΩΝ (με κριτές)

[P01] Agata Savary, Daniel Zeman, Verginica Barbu Mititelu, Anabela Barreiro, Olesea Caftanatot, Marie-Catherine de Marneffe, Kaja Dobrovoljc, Gülşen Eryiğit, **Voula Giouli**, Bruno Guillaume, Stella Markantonatou, Nurit Melnik, Joakim Nivre, Atul Kr. Ojha, Carlos Ramisch, Abigail Walsh, Beata Wójtowicz, and Alina Wróblewska. 2024. [UniDive: A COST Action on Universality, Diversity and Idiosyncrasy in Language Technology](#). In *Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages @ LREC-COLING 2024*, pages 372–382, Torino, Italia. ELRA and ICCL.

[P02] Verginica Barbu Mititelu, Voula Giouli, Kilian Evang, Daniel Zeman, Petya Osenova, Carole Tiberius, Simon Krek, Stella Markantonatou, Ivelina Stoyanova, Ranka Stanković, and Christian Chiarcos. 2024. [Multiword Expressions between the Corpus and the Lexicon: Universality, Idiosyncrasy, and the Lexicon-Corpus Interface](#). In *Proceedings of the Joint Workshop on Multiword Expressions and Universal Dependencies (MWE-UD) @ LREC-COLING 2024*, pages 147–153, Torino, Italia. ELRA and ICCL.

[P03] Savary, A., Ben Khelil. Ch., Ramisch, C., **Giouli**, V., Barbu, Mititelu, et al., (2023). PARSEME Corpus release 1.3. In *Proceedings of the 19th Workshop of Multi-word expressions (MWE 2023)*. Collocated with the European Association of Computational Linguistics (EACL 2023), Dubrovnik, Croatia. May 6, 2023.

[P04] Bikakis, N., Giannopoulos, G., Sidiropoulos, N., Flouda, Ch., Doupas, A., **Giouli**, V., Karioris, P., Botini, P., Vacalopoulou, A., and Stainhaouer, G., (2023). Exposing Geospatial Cultural Heritage Content in Map-based Applications. In *Proceedings of BigVis Workshop, collocated with the 26th International Conference on Extending Database Technology jointly organised with the 26th International Conference on Database Theory (EDBT/ICDT joint 2023 Conference), March 28 – 31, 2023, Ioannina, Greece*.

[P05] **Giouli**, V., Vacalopoulou, A., Sidiropoulos, N., Flouda, C., Doupas, A., & Stainhaouer, G. (2022). From Mythos to Logos: a bilingual thesaurus tailored to meet users' needs within the ecosystem of Cultural Tourism. In *Proceedings of the XX EURALEX International Congress*, July 2022, Marseille, France.

[P06] **Giouli**, V., Vacalopoulou, A., Sidiropoulos, N., Flouda, C., Doupas, A., Giannopoulos, G., Bikakis, N., Kaffes, V., & Stainhaouer, G. (2022). Placing multi-modal, and multi-lingual Data in the Humanities

Domain on the Map: The Mythotopia Geo-tagged Corpus. In Proceedings of the 13th chapter of the *International Language Resources and Evaluation Conference*, 20-25 June 2022, Marseille, France.

- [P07] Pilitsidou, V., and **Giouli**, V. 2021. Frame Semantics in the specialized domain of Finance: building a termbase to aid Translation. In *Proceedings of the XIX EURALEX International Congress*, 8-12 September 2020, Alexandroupolis, Greece.
- [P08] **Giouli**, V., and Sidiropoulos, N. 2021. Making dictionaries visible, accessible, and reusable: the case of the Greek Conceptual dictionary API. In *Proceedings of the XIX EURALEX International Congress*, 8-12 September 2020, Alexandroupolis, Greece.
- [P09] Ramisch, C., Savary, A., Guillaume, B., Waszczuk, J., Candito, M., Vaidya, A., Barbu Mititelu, V., Bhatia, A., Inurrieta, U., **Giouli**, V., Gungor, T., Jiang, M., Lichte, T., Liebeskind, C., Monti, J., Ramisch, R., Stymne, S., Walsh, A., Xu, H. 2020. Edition 1.2 of the PARSEME Shared Task on Semi-supervised Identification of Verbal Multiword Expressions. In *Proceedings of the Joint Workshop on Multiword Expressions and Electronic Lexicons (MWE-LEX 2020)*, at COLING 2020 (Barcelona, Spain online), December 13, 2020.
- [P10] **Giouli**, V., Pilitsidou, V., and Christopoulos, H. 2020. Greek within the Global FrameNet Initiative: Challenges and Conclusions so far. Στο *Proceedings of the International FrameNet Workshop 2020: Towards a Global, Multilingual FrameNet*, pp. 48 – 55. Language Resources and Evaluation Conference (LREC 2020), 11 – 16 May, Marseille, France.
- [P11] **Giouli**, V., Fotopoulou, A., and Foufi, V. 2019. Annotating VMWEs in running text: a piece of cake or looking for a needle in a haystack? In *Proceedings of the 13th International conference on Greek Linguistics*, pp. 125 - 134. 7-9 September 2017. Westminster University, London, UK.
- [P12] Ramisch, C., Cordeiro, S., Savary, A., Vincze, V., Mititelu, V., Bhatia, A., Buljan, M., Candito, M., Gantar, P., **Giouli**, V., GÜNGÖR, T., Hawwari, A., Iñurrieta, U., Kovalevskaitė, J., Krek, S., Lichte, T., Liebeskind, C., Monti, J., Parra Escartín, C., Qasemizadeh, B., Ramisch, R., Schneider, N., Stoyanova, I., Vaidya, A., Walsh, A. 2018. Edition 1.1 of the PARSEME Shared Task on Automatic Identification of Verbal Multiword Expressions. In *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*, pp. 222 - 240.
- [P13] Savary, A. Ramisch, C., Cordeiro, S., Sangati, F., Vincze, V., QasemiZadeh, B., Candito, M., Cap, F., **Giouli**, V., Stoyanova, I., and Doucet, A. 2017. The PARSEME Shared Task on Automatic Identification of Verbal Multiword Expression. In *Proceedings of the 13th Workshop on Multiword Expressions at the European Chapter of the Association for Computational Linguistics (EACL) Conference*, pp. 31-47, Valencia, Spain.
- [P14] Φωτοπούλου, Α., και **Γιούλη**, Β. 2017. Από την Έκφραση στο Πολύτροπο: σχεδιασμός και οργάνωση ενός εννοιολογικού λεξικού. Στο *Proceedings of the ICGL12*, Vol. 1, pp. 327-339, Berlin, Edition Romiosini/CeMoG.
- [P15] Dorn, A., Villalva, A., **Giouli**, V., Blanck, W., Kovalenco, K., and Wandl-Vogt, E. 2016. Displaying Language Diversity on the European Dictionary Portal – COST-ENeL-Case Study on Colours and Emotions and their cultural differences. In *Proceedings of the XVII EURALEX International Congress: Lexicography and Linguistic Diversity*, pp. 237 – 243. Tbilisi, Georgia, 6-10 September 2016. Tbilisi University Press.
- [P16] **Giouli**, V. 2015. Populating a lexicon with Multiword Expressions in view of Machine Translation. In *Proceedings of the 2nd Workshop on Multi-word Units in Machine Translation and Translation Technology (MUMTTT2015)*, pp. 62 - 67. Malaga, Spain, 1-2 July 2015.

- [P17] Fotopoulou, A. and Giouli, V. 2015. MWEs: Support/light verb constructions vs frozen constructions in Modern Greek and French. In *Proceedings of the 2nd Workshop on Multi-word Units in Machine Translation and Translation Technology (MUMTTT2015)*, pp. 68-73. Malaga, Spain, 1-2 July 2015.
- [P18] Giouli, V. and Fotopoulou, A. 2014. Linguistically motivated Language Resources for Sentiment Analysis. In *Proceedings of the Workshop for Lexical and Grammatical Resources for Language Processing*, pp. 39-45. *25th International Conference on Computational Linguistics (COLING 2014)*, Dublin, Ireland, 24 August 2014.
- [P19] Fotopoulou, A. Markantonatou, S. and Giouli, V. 2014. Encoding MWEs in a conceptual lexicon. In *Proceedings of the 10th Workshop on Multiword Expressions (MWE 2014)*, pp. 43-47. Gothenburg, Sweden. 26–27 April 2014.
- [P20] Giouli, V. and Fotopoulou, A. 2014. Developing Language Resources for Sentiment Analysis in Greek. In G. Kotzoglou et al. (eds), *Selected Papers of the 11th International Conference on Greek Linguistics*, pp. 501-511. Rhodes: University of the Aegean.
- [P21] Giouli, V., and A. Fotopoulou. 2012a. A Local Grammar of verb predicates denoting Emotion in Greek: A Lexicon-Grammar account. In Radimsky, J. (Ed.) *Proceedings of the 31st International Conference on Lexis and Grammar 2012*. pp. 70-74. Nové Hradý, Czech Republic, 19- 22 September 2012. 978-80-7394-409-4. hal-03321093.
- [P22] Giouli, V., and A. Fotopoulou. 2012b. Emotion verbs in Greek. From Lexicon-Grammar tables to multi-purpose syntactic and semantic lexica. In *Proceedings of the XV Euralex International Congress (EURALEX 2012)*, pp. 485-492. Oslo, Norway. August 2012.
- [P23] Vacalopoulou, A., Giouli, V., Efthimiou, E., Giagkou, M. 2012. Bridging the gap between disconnected languages: the eMiLang multi-lingual database. In *Proceedings of the Workshop on Creating Cross-language Resources for Disconnected Languages and Styles (CREDISLAS, 2012)*. Istanbul, Turkey, 27 May 2012.
- [P24] Mouka, E., V. Giouli, A. Fotopoulou and I.E. Saridakis. 2012. Opinion and emotion in movies: a modular perspective to annotation. In *Proceedings of the 4th International Workshop on Corpora for Research on Emotion, Sentiment & Social Signals (ES³ 2012)*, pp. 104 – 109. Istanbul, Turkey, 26 May 2012.
- [P25] Γιούλη, Β., και Α. Φωτοπούλου. 2012. Σημασιολογικό πεδίο των «συναισθημάτων»: ταξινόμηση των ρημάτων της ελληνικής που δηλώνουν «συναίσθημα». *10th International Conference in Greek Linguistics*, Κομοτηνή, 1-4 Σεπτεμβρίου 2011.
- [P26] Vacalopoulou, A., Giouli, V., Giagkou, M., and Efthimiou, E. 2011. Online Dictionaries for immigrants in Greece: Overcoming the Communication Barriers. In *Proceedings of the 2nd Conference “Electronic Lexicography in the 21st century: new Applications for New users” (eLEX2011)*, Bled, Slovenia, 10-12 November 2011.
- [P27] Giouli, V., Marzelou, E., Prokopidis, P., and Zourari, M. 2009. Language Technologies for Processing Greek textual Cultural Heritage data. In *Proceedings of the 9th International Conference on Greek Linguistics, ICGL9*, Chicago, 29 - 31 October 2009
- [P28] Giouli, V., Glaros, N. Simov, K., and Osenova, P. 2009. A web-enabled and speech-enhanced parallel corpus of Greek - Bulgarian cultural texts. In *Proceedings of the EACL 2009 Workshop on Language*

- [P29] Gavrilidou, M., Giouli, V., Lambropoulou, P. 2008. The Greek High School Dictionary: Description and issues. In *Proceedings of the XIII Euralex International Congress (EURALEX 2008)*. July 2008, Barcelona, Spain.
- [P30] Marzelou, E., Zourari, M., Giouli, V., Piperidis, S., 2008. Building a Greek corpus of Textual Entailment. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*. Marrakech, Morocco.
- [P31] Γαβριηλίδου, Μ., Λαμπροπούλου, Π., Γιούλη, Β., 2007. Το Ερμηνευτικό Λεξικό για το Γυμνάσιο, *8th International Conference on Greek Linguistics*, Ιωάννινα 31/8 – 2/9/2007.
- [P32] Giouli, V., Konstandinidis, A., Desypri, E., Papageorgiou, H. 2006. Multi-domain Multi-lingual Named Entity Recognition: Revisiting & Grounding the resources issue. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, Genoa, Italy.
- [P33] Gavrilidou, M., Labropoulou, P., Piperidis, S., Giouli, V., Calzolari, N., Monachini, M., Soria, C., Choukri, K., 2006. Language Resources Production Models: the case of the INTERA Multilingual Corpus and Terminology. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, Genoa, Italy.
- [P34] Gavrilidou, M., Labropoulou, P., Desypri, E., Giouli, V., Antonopoulos, V., Piperidis, S., 2004. Building Parallel Corpora for eContent Professionals. In *Proceedings of the COLING Workshop on Multilingual Linguistic Resources (MLR2004)*, Geneva, Switzerland.
- [P35] Papageorgiou, H., Prokopidis, P., Giouli, V., Demiros, I., Konstantinidis, A., Piperidis, S., 2003. Human Language Technology and Processing of Greek Language. In *Proceedings of the 6th International conference of Greek linguistics (ICGL 2003)*, Rethymno, Greece.
- [P36] Papageorgiou, H., Prokopidis, P., Giouli, V., Demiros, I., Konstantinidis, A., Piperidis, S., 2002. Multi-level, XML - based Corpus Annotation. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, Las Palmas, Spain.
- [P37] Γιούλη, Β., Δεμοίρος, Ι., Αντωνόπουλος, Β., Παπαγεωργίου, Χ., Πιπερίδης, Σ., Αρβανίτης, Π., Παπαναστασίου, Ε., Κακλαμανίδου, Δ., Χιδίρογλου, Ά., Αλεξανδροπούλου, Α., Κυλινηρέα, Α., Κοπιδάκης, Μ., 2002. "οικΟΝΟΜία": Ένα Σώμα Σχολιασμένων Οικονομικών Κειμένων", Στο *Μελέτες για την ελληνική γλώσσα. Πρακτικά της 22ης ετήσιας συνάντησης του Τομέα Γλωσσολογίας της Φιλοσοφικής Σχολής του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης (27-29 Απριλίου 2001)*, τόμ. 2, σελ. 20-31.
- [P38] Γιούλη, Β., Μπούτσης, Σ., Δεμοίρος, Ι., Αντωνόπουλος, Β., Παπαγεωργίου, Χ., Πιπερίδης, Σ., Αρβανίτης, Π., Χιδίρογλου, Ά., Κοπιδάκης, Μ. 2001. Αυτόματη Αναγνώριση Ονοματικών Οντοτήτων για Εξαγωγή και Ανάκτηση Πληροφοριών. Στο *Μελέτες για την ελληνική γλώσσα. Πρακτικά της 21ης ετήσιας συνάντησης του Τομέα Γλωσσολογίας της Φιλοσοφικής Σχολής του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης (12-14 Μαΐου 2000)*, σελ. 133-143.
- [P39] Μπούτσης, Σ., Προκοπίδης, Π., Γιούλη, Β., Πιπερίδης, Σ. 2001. Επιφανειακή Συντακτική Ανάλυση Ελευθέρου Κειμένου. Στο *Μελέτες για την Ελληνική Γλώσσα, Στο Μελέτες για την ελληνική γλώσσα. Πρακτικά της 21ης ετήσιας συνάντησης του Τομέα Γλωσσολογίας της Φιλοσοφικής Σχολής του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης (12-14 Μαΐου 2000)*, σελ. 502-513.

- [P40] Demiros, I., Boutsis, I., Giouli, V., Liakata, M., Papageorgiou, H., Piperidis, S. 2000. Named Entity Recognition in Greek Texts. In *Proceedings of the Second International Conference on Language Resources and Evaluation*, Athens, Greece.
- [P41] Boutsis, S., Prokopidis, P., Giouli, V., Piperidis, S. 2000. A robust Parser for unrestricted Greek Text. In *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC2000)*, pp. 467-474. Athens, Greece.
- [P42] Papageorgiou, H., Prokopidis, P., Giouli, V., Piperidis, S. 2000. A unified tagging Architecture and its Application to Greek. In *Proceedings of the Second International Conference on Language Resources and Evaluation*, Athens, Greece.
- [P43] Κουτσογεωργοπούλου, Ε., Γιούλη, Π., Δερματάς, Ε., Κοκκινάκης, Γ. 1999. Ανάπτυξη Ηλεκτρονικών Λεξικών της ελληνικής γλώσσας. Στο *Μελέτες για την ελληνική γλώσσα. Πρακτικά της 19ης ετήσιας συνάντησης του Τομέα Γλωσσολογίας της Φιλοσοφικής Σχολής του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης* (23-25 Απριλίου 1998), σελ. 174-188.
- [P44] Γιούλη, Π. 1995. Η Εξέλιξη της Δοτικής. Στο *Μελέτες για την ελληνική γλώσσα. Πρακτικά της 15^{ης} ετήσιας συνάντησης του Τομέα Γλωσσολογίας της Φιλοσοφικής Σχολής του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης*, σελ. 135-146.

Παρουσίαση σε συνέδρια (χωρίς άρθρο)

Giouli, V. (2023). Morphology in a computational semantic lexical resource: theoretical, lexicographic, and encoding considerations. Accepted for presentation at the 16th International Conference on Greek Linguistics, 14-17 December 2023, Thessaloniki, Greece.

Christopoulos, H., and Giouli, V. (2023). Encoding COGNITION, COMMUNICATION and STANCE predicates in a semantic lexicon of Modern Greek: a frame-based account. Accepted for presentation at the 16th International Conference on Greek Linguistics, 14-17 December 2023, Thessaloniki, Greece.

Giouli, V. 2023. (2023). Building a Frame-based lexical resource for Modern Greek: linguistic, lexicographic, and cognitive considerations. Paper presented at the *43rd Annual Meeting of the Department of Linguistics, Aristotle University of Thessaloniki*, 5-7 May, 2023.

Giouli, V., Pilitsidou, V., and Christopoulos, H. (2023). The Greek FrameNet project: populating and interlinking a lexical resource. Paper accepted at the *International FrameNet Workshop 2023 (IFNW 2023): Cognitively Grounded Approaches to Applied Language Description*, collocated with the *16th International Cognitive Linguistics Conference (ICLC 2023)*, Düsseldorf, Germany, 7 – 11 August 2023.

Vacalopoulou, A., Sidiropoulos, N., Flouda, C., Doupas, A., Giouli, V., Karioris, P., Botini, P., Giannopoulos, G., Bikakis, N., and Stainhaouer, G. (2023). In view of reusable representation of cultural content: the case of a multilingual and multimodal corpus in the Greek context. *HUGOD Conference: Humanities Going Digital: Teaching, training & research experiences*, February 2-3, 2023.

Giouli, V., Vacalopoulou, A., Giannopoulos, G., Doupas, A., Karioris, P., Michailidou, N., Mourthos, I., Bikakis, N., Botini, P., Sarafopoulou, D., Sidiropoulos, N., Steinhauer, G., Tsiafaki, D., and Flouda, Ch. (2022). Language Resources for human-centric NLP: the MYTHOTOPIA geo-tagged dataset. Poster presentation at the *1st Workshop on Human-Centric Sciences and Technologies*, Xanthi, Greece, 20 October.

Mouka, E., I. E. Saridakis, A. Fotopoulou, and V. Giouli. 2013. Annotating racism: A corpus-based study of cross-linguistic racist discourse: annotation and analysis. *The 7th Congress of the European Society for Translation Studies (7th EST) Translation Studies: Centres and Peripheries*. Germersheim, Germany, 29 – 31 August 2013.

ΠΡΑΚΤΙΚΑ ΣΥΝΕΔΡΙΩΝ - ΕΠΙΣΤΗΜΟΝΙΚΩΝ ΗΜΕΡΙΔΩΝ

Archna Bhatia, Gosse Bouma, A. Seza Dogruoz, Kilian Evang, Marcos Garcia, Voula Giouli, Lifeng Han, Joakim Nivre, and Alexandre Rademaker. 2024. [*Proceedings of the Joint Workshop on Multiword Expressions and Universal Dependencies \(MWE-UD\) @ LREC-COLING 2024*](#). ELRA and ICCL, Torino, Italia, edition.

Archna Bhatia, Kilian Evang, Marcos Garcia, Voula Giouli, Lifeng Han, and Shiva Taslimipoor. 2023. [*Proceedings of the 19th Workshop on Multiword Expressions \(MWE 2023\)*](#). Association for Computational Linguistics, Dubrovnik, Croatia, edition.

ΤΕΧΝΙΚΕΣ ΑΝΑΦΟΡΕΣ (επιλεγμένες)

[TA01] Candito, M., Cap, F., Cordeiro, S., Foufi, V., Gantar, P., Giouli, V., Herrero, C., Ionescu, M., Mititelu, V., Monti, J., Nivre, J., Onofrei, M., Parra Escartín, C., Sailer, M., Ramisch, C., Rizea, M. M., Savary, A., Stonayova, I., Stymne, S., and Vincze, V. 2017. Annotation guidelines - PARSEME shared task on automatic identification of verbal MWEs. <http://hdl.handle.net/11574/170898>.

[TA02] Giouli, Voula, Radu Ion, Gasper Koren, Madara Mieriņa, Serge Sharoff, Inguna Skadiņa, Marco Tadic, Gregor Thurmair. 2010. Test corpora of comparable texts to evaluate comparability metrics. ACCURAT Technical Report.

[TA03] Giouli, V. and Stelios Piperidis. 2002. *Corpora and HLT. Current trends in corpus processing and annotation*. IST BALRIC-RIC Technical Report.

VIII. ΛΟΙΠΑ ΕΡΓΑ

ΛΕΞΙΚΑ (έντυπα ή/και ηλεκτρονικά)

Εννοιολογικό Λεξικό της Νέας Ελληνικής. Ινστιτούτο Επεξεργασίας του Λόγου, Ερευνητικό Κέντρο ΑΘΗΝΑ. API.

Αναστασιάδου – Συμεωνίδου, Ά., Β. Γιούλη, Ε. Κουτσογεωργοπούλου και Γ. Κοκκινάκης. 2008. *ΚΟΡΑΗΣ. Έντυπο και Ηλεκτρονικό Ελληνο-αγγλικό Λεξικό*. Εκδόσεις Πανεπιστημίου Πατρών, 2008. ISBN: 9606843181.

Ερμηνευτικό Λεξικό για το Γυμνάσιο, Οργανισμός Εκδόσεως Διδακτικών Βιβλίων, 2008.

Electronic Dictionary of Pronunciation and Usage of Modern Greek for Foreigners. Brasshouse Language Centre. Birmingham, UK, 1998.

COLLINS. Greek-English, English-Greek Dictionary. HarperCollins Publishers, Ltd, UK.

The Graecanic Lexicon. Τρίγλωσσο ηλεκτρονικό λεξικό κατωιταλικής - νέας ελληνικής – ιταλικής. Εργαστήριο Ενσύρματης Τηλεπικοινωνίας, Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών, Πανεπιστήμιο Πάτρας, 1998.

Λεξικά του έργου eMilang (9 δίγλωσσα λεξικά από τα Ελληνικά προς καθεμία από τις γλώσσες Αλβανικά, Βουλγαρικά, Ρουμανικά, Ρωσικά, Σερβικά, Πολωνικά, Κινέζικα-Μανδαρινικά, Αραβικά και Αγγλικά).
<http://www.e-milang.gr/>

ΓΛΩΣΣΙΚΟΙ ΠΟΡΟΙ (ενδεικτικά)

ConceptNet-el. Σημασιολογικό Λεξικό της Νέας Ελληνικής. Institute for Language and Speech Processing - Athena Research Center (2021, July 08). ConceptNet-el. Version 2.0.0. [Dataset (Lexical/Conceptual Resource)]. CLARIN:EL. <http://hdl.handle.net/11500/CLARIN-EL-0000-0000-765A-7>

Σημασιολογικό Λεξικό της Νέας Ελληνικής στο πρότυπο της Σημασιολογίας Πλαισίων (FrameNet-EL v.1.0). Διαθέσιμο μέσω της υποδομής ΑΠΟΛΛΩΝΙΣ και του ιστότοπου του έργου Global FrameNet.

Πίνακες του Λεξικού-Γραμματικής. Υπολογιστικό λεξικό και υπολογιστική γραμματική των ρημάτων ψυχικού πάθους της Ελληνικής. Διαθέσιμο μέσω της ερευνητικής Υποδομής CLARIN-ΑΠΟΛΛΩΝΙΣ. <http://hdl.handle.net/11500/CLARIN-EL-0000-0000-75F7-6>.

Λεξικογραφική Βάση Δεδομένων ονοματικών και ρηματικών πολυλεκτικών εκφράσεων της Νέας Ελληνικής PolyLEX. Παραδοτέο έργου *Υπολογιστικές Επιστήμες και Τεχνολογίες Δεδομένων, Περιεχομένου και Αλληλεπίδρασης* (ΟΠΣ 5002437). Διαθέσιμη μέσω της υποδομής ΑΠΟΛΛΩΝΙΣ.

PARSEME-ST-v1.3 Greek corpus of Verbal Multiword Expressions. Το ελληνικό υπο-σώμα κειμένων (ELL section) του the PARSEME VMWEs corpus v.1.3 αποτελείται από κείμενα επισημειωμένα με ρηματικές πολυλεκτικές εκφράσεις (Verbal Multi-word expressions, VMWEs) στη βάση σχετικών προδιαγραφών. Στόχος είναι η εκπαίδευση και αξιολόγηση υπολογιστικών εργαλείων αναγνώρισης ρηματικών πολυλεκτικών. Η παρούσα έκδοση περιλαμβάνει διορθώσεις της προηγούμενης έκδοσης (v.1.2), κυρίως στο επίπεδο της μορφοσυντακτικής ανάλυσης και λημματοποίησης, αλλά και στο επίπεδο των πολυλεκτικών, ενώ έχει εμπλουτιστεί και με νέα κείμενα. <http://hdl.handle.net/11372/LRT-5124>.

PARSEME-ST-v1.2 Greek corpus of Verbal Multiword Expressions. Το ελληνικό υπο-σώμα κειμένων (ELL section) δημιουργήθηκε στο πλαίσιο του διαγωνισμού PARSEME Shared Task on the identification of VMWEs (για το έτος 2020), ως τμήμα του ευρύτερου PARSEME-ST-v1.2 corpus. Στο σώμα κειμένων έχουν αναγνωρισθεί και επισημειωθεί με χειροκίνητο τρόπο ρηματικές πολυλεκτικές εκφράσεις (Verbal Multi-word expressions, VMWEs) στη βάση προδιαγραφών επισημείωσης (έκδοση 1.2), με στόχο την εκπαίδευση και αξιολόγηση υπολογιστικών εργαλείων αναγνώρισης πολυλεκτικών εκφράσεων. <http://hdl.handle.net/11234/1-3416>.

PARSEME-ST-v.1.2 Greek morpho-syntactically annotated corpus for the PARSEME Shared Task on Semi-Supervised Identification of Verbal Multiword Expressions (edition 1.2). Ελληνικό σώμα κειμένων το οποίο δημιουργήθηκε στο πλαίσιο του διαγωνισμού PARSEME Shared Task on the identification of VMWEs (για το έτος 2020), ως τμήμα του ευρύτερου PARSEME-ST-v1.2 corpus. Φέρει μορφο-συντακτική επισημείωση (λήμμα, γραμματική κατηγορία, μορφοσυντακτικά χαρακτηριστικά και συντακτικές εξαρτήσεις) σε μορφή CONLL-U (<https://universaldependencies.org/format.html>). Αποτελεί πρόσθετο "ακατέργαστο" υλικό που δόθηκε στους συμμετέχοντες του διαγωνισμού, με στόχο τη μη επιβλεπόμενη μάθηση για τον εντοπισμό και την αναγνώριση ρηματικών πολυλεκτικών εκφράσεων που δεν απαντώνται στα επισημειωμένα κείμενα (unseen VMWEs). <http://hdl.handle.net/11234/1-3416>

PARSEME-ST-v1.1 Greek corpus of Verbal Multiword Expressions. Το ελληνικό υπο-σώμα κειμένων (ELL section) δημιουργήθηκε στο πλαίσιο του διαγωνισμού PARSEME Shared Task on the identification of VMWEs (για το έτος 2019), ως τμήμα του ευρύτερου PARSEME Shared Task on the identification of VMWEs corpus v.1.1. Το σώμα κειμένων φέρει επισημείωση ρηματικών πολυλεκτικών εκφράσεων, η οποία έχει πραγματοποιηθεί με τρόπο χειροκίνητο στη βάση σχετικών προδιαγραφών (επικαιροποιημένη έκδοση 1.1), με στόχο την εκπαίδευση και αξιολόγηση εργαλείων αναγνώρισης και

επισημείωσης ρηματικών πολυλεκτικών. Διαθέσιμο στο <http://hdl.grnet.gr/11500/ATHENA-0000-0000-5E7F-A>.

PARSEME-ST-v1.0. Greek Corpus of Verbal Multiword Expressions. Το ελληνικό υπο-σώμα κειμένων (ELL section) δημιουργήθηκε στο πλαίσιο του διαγωνισμού PARSEME Shared Task on the identification of VMWEs, ως τμήμα του ευρύτερου PARSEME-ST-v1.0 corpus. Στο σώμα κειμένων έχουν αναγνωρισθεί και επισημειωθεί με χειροκίνητο τρόπο ρηματικές πολυλεκτικές εκφράσεις (Verbal Multi-word expressions, VMWEs) με βάση τις προδιαγραφές επισημείωσης πολυλεκτικών εκφράσεων (έκδοση 1.0), με στόχο την εκπαίδευση και αξιολόγηση υπολογιστικών εργαλείων αναγνώρισης πολυλεκτικών εκφράσεων. <http://hdl.grnet.gr/11500/ATHENA-0000-0000-5E80-6>.

Εννοιολογικό Λεξικό ΠΟΛΥΤΡΟΠΟΝ. Ινστιτούτο Επεξεργασίας του Λόγου του Ερευνητικού Κέντρου ΑΘΗΝΑ. <http://hdl.grnet.gr/11500/ATHENA-0000-0000-2862-5>.

Ελληνικό Σώμα Κειμενικής Συνεπαγωγής GTEC. Σώμα κειμένων της Ελληνικής το οποίο αποτελείται από ζεύγη προτάσεων, καθένα από τα οποία φέρει επισημείωση ανάλογα με το αν το σημασιολογικό περιεχόμενο της μίας πρότασης (πρότασης ελέγχου, Test sentence) συνάγεται λογικά από το σημασιολογικό περιεχόμενο της άλλης (Υπόθεση, Hypothesis Sentence). <http://hdl.grnet.gr/11500/ATHENA-0000-0000-23DB-2>.

Savary, A., S. R. Cordeiro, T. Lichte, C. Ramisch, U. Iñurrieta, and V. Giouli. 2019. Multilingual corpus of literal occurrences of multiword expressions, Digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. <http://hdl.handle.net/11372/LRT-2966>.

ΕΡΓΑΛΕΙΑ ΕΠΕΞΕΡΓΑΣΙΑΣ ΦΥΣΙΚΗΣ ΓΛΩΣΣΑΣ (ενδεικτικά)

Υπήρξα βασικό στέλεχος της ομάδας ανάπτυξης των εργαλείων επεξεργασίας της Ελληνικής του IEL:

Εργαλείο Ληματοποίησης κειμένων της Ελληνικής (ilsp_lemmatiser): δημιουργία σώματος κειμένων, κατάρτιση κανόνων, αποσφαλμάτωση εργαλείου

Εργαλείο Μορφοσυντακτικής Ανάλυσης και Ληματοποίησης (FBT part-of-speech tagger): συμμετοχή στην κατάρτιση των προδιαγραφών και την οριστικοποίηση του συνόλου χαρακτηριστικών (tagset), επισημείωση σώματος κειμένων, αποσφαλμάτωση εργαλείου

Εργαλείο Συντακτικής Ανάλυσης της Ελληνικής (ilsp_chunker): συγκρότηση και επισημείωση σώματος κειμένων, κατάρτιση προδιαγραφών, συγγραφή των γραμματικών κανόνων, αποσφαλμάτωση εργαλείου.

Εργαλείο Αναγνώρισης Ονοματικών Οντοτήτων της Ελληνικής (βασισμένο σε κανόνες) (ilsp_nerc): συγκρότηση και επισημείωση σώματος κειμένων, κατάρτιση προδιαγραφών, κατάρτιση βοηθητικών πόρων (gazeteers), συγγραφή των γραμματικών κανόνων, αποσφαλμάτωση εργαλείου.

Εργαλείο Αναγνώρισης Ονοματικών Οντοτήτων σε ελληνικά και αγγλικά κείμενα (Maximum Entropy) (ilsp_nes): συγκρότηση και επισημείωση σώματος κειμένων, κατάρτιση προδιαγραφών και του νέου διευρυμένου και αρθρωτού σχήματος επισημείωσης, συγγραφή των γραμματικών κανόνων.

IX. ΣΕΜΙΝΑΡΙΑ ΕΞΕΙΔΙΚΕΥΣΗΣ

2020 Working with Corpora and Digital Tools. Facultad de Filología, Universidad Complutense, Madrid, Spain. 22- 28 Οκτωβρίου 2020 (διαδικτυακά).

- 2020 Corpus Linguistics LAEL Webinars. Graduate Program in Applied Linguistics, Pontifical Catholic University of Sao Paulo, Brazil. Απρίλιος - Σεπτέμβριος 2020 (διαδικτυακά).
- 2017 Annotation of Discourse Relational Devices (DRDs): Multilingual and Multimodal Challenges. Facultad de Filología, Universidad Complutense, Madrid, Spain. 12-14 June 2017.
- 2010 Σεμινάριο *Μετάφρασης Παιδικής Λογοτεχνίας*. Κέντρο Μεταφραστικών Σπουδών meta|φραση.
- 1994 *Courses for English as a Foreign Language (EFL) teachers*, TESOL Greece.
- 1993 *Εξειδίκευση στην Ηλεκτρονική Λεξικογραφία*. Εργαστήριο Ενσύρματης Επικοινωνίας, Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών, Πανεπιστήμιο Πατρών. Σε συνεργασία με τον εκδοτικό οίκο Collins και το λεξικογραφικό κέντρο Cobuild.
- 1991 Σεμιναριακό μάθημα *Εφαρμοσμένη Γλωσσολογία: Διδασκαλία της Γλώσσας στην Εκπαίδευση*. Εθνικό & Καποδιστριακό Πανεπιστήμιο Αθηνών, Φιλοσοφική Σχολή, Τμήμα Φιλολογίας, Κατεύθυνση Γλωσσολογίας. Εκπόνηση σπουδαστικής εργασίας με τίτλο: *Διδακτική της Γλώσσας στο Δημοτικό μέσα από τις ασκήσεις των σχολικών εγχειριδίων*.

ΞΕΝΕΣ ΓΛΩΣΣΕΣ

Αγγλικά: Επίπεδο C2. Πτυχία *First Certificate in English* University of Cambridge, και *Certificate of Proficiency in English*, University of Cambridge. Επάρκεια διδασκαλίας.

Γαλλικά: Επίπεδο B2. *Certificat de Langue Francaise*.

Γερμανικά: Αρχάρια.

ΓΝΩΣΗ Η/Υ

Λειτουργικά Συστήματα: Windows, MacOS, UNIX, LINUX.

Γλώσσες Προγραμματισμού: Prolog, Perl, Python.

Εργαλεία Επισημείωσης Κειμένων: (Inception, WebAnno, TrEd, GNOME, BRAT, FLAT, INCEPTION, Ellogon, ELAN).

Διάφορα: MS Office/OpenOffice, MS Project, LTC Organiser, EMACS, LATEX, HTML, XML, Overleaf.

Εργαλεία Υποβοηθούμενης Μετάφρασης (Computer-Assisted Translation Tools) και Μεταφραστικών Μνημών (Trados, DejaVu, OmegaT)

ΠΡΟΣΩΠΙΚΑ ΕΝΔΙΑΦΕΡΟΝΤΑ

Μουσική. Σπουδές κλασικής κιθάρας με ολοκλήρωση του κύκλου σπουδών μέχρι και το επίπεδο της Β' Ανωτέρας, μαθήματα σολφέζ και αρμονίας (Πατραϊκό Ωδείο, Πάτρα).

Παιδική Λογοτεχνία, Μετάφραση Παιδικής Λογοτεχνίας.

Διάβασμα, θέατρο, Κινηματογράφος, Μαγειρική, κηπουρική.

Τρέξιμο, κολύμβηση, τένις.