

22 Λεξιλογικές και κοινωνιοσημειωτικές πτυχές της γλωσσικής ποικιλίας της Κωνσταντινούπολης: μελέτη βασισμένη σε σώμα κειμένων

Νίκη Χρ. Τρύφωνα

ΠΕΡΙΛΗΨΗ

Αντικείμενο της μελέτης είναι η συστηματοποίηση και κριτική ερμηνεία, υπό κοινωνιογλωσσολογικό πρίσμα, πτυχών της χρήσης της γλώσσας από τα μέλη της ελληνορθόδοξης μειονότητας της Κωνσταντινούπολης – γλωσσικής ποικιλίας που είναι ευρύτερα γνωστή ως *Πολίτικα* και που έως σήμερα έχει τύχει ελάχιστης ερευνητικής προσοχής. Η μελέτη εστιάζει περαιτέρω στη συστηματική αντιπαραβολική διερεύνηση των (εστιακών) καταγραφόμενων γλωσσικών επιλογών, κυρίως στο λεξιλογικό επίπεδο, με την Κοινή Νεοελληνική, με έμφαση στον λεξιλογικό δανεισμό και την κοινωνική σημείωση. Η έρευνα ακολουθεί το γενικό ευρετικό και ερευνητικό μοντέλο της Γλωσσολογίας Σωμάτων Κειμένων, καθοδηγούμενη από αυθεντικά κειμενικά δεδομένα, στη δε κριτική προσέγγισή της συνδυάζει προτάγματα της υποβοηθούμενης από σώματα κειμένων ανάλυσης λόγου (CADS). Το μείζον μέρος των ευρημάτων συγκεντρώνεται σε ψηφιακό (ορολογικό) πόρο, δυνάμενο να επεκταθεί οριζόντια, με διεύρυνση (χρονική, ποσοτική) του σώματος κειμένων και κάθετα, με συμπερίληψη περισσότερων επιπέδων χαρακτηρισμού (επισημείωσης).

Lexical and socio-semiotic aspects of the linguistic variety of Istanbul: a corpus-based study

Nicky Chr. Tryfona

ABSTRACT

The aim of this study is to systematise and critically interpret, from a sociolinguistic perspective, aspects of language use by members of the Greek Orthodox minority of Istanbul – a linguistic variety that is more widely known as *Politika* and that has received little research attention to date. The study further focuses on the systematic contrastive investigation of the (focal) identified linguistic choices, mainly at the lexical level, with Modern Greek, with an emphasis on lexical borrowing and social signification. The research follows the general heuristic and research model of Corpus Linguistics, guided by authentic textual data, and in its critical approach combines principles of Corpus-Assisted Discourse Studies (CADS). The majority of the findings are concentrated in a digital (terminological) resource, of which the underlying language resource (corpus) is expandable horizontally, i.e. longitudinally and in size, and vertically, by including more levels of discourse annotation.

0 Εισαγωγή και υπόβαθρο μελέτης

Η παρούσα μελέτη είναι διεπιστημονική, στην ευρετική μεθοδολογία της και στην ερμηνεία των ευρημάτων της. Βασίζεται κυρίως στα επιστημονικά υποδείγματα της Γλωσσολογίας Σωμάτων Κειμένων (ΓΣΚ), αξιοποιώντας στοιχεία της ηλεκτρονικής λεξικογραφίας και της (κριτικής), υποβοηθούμενης από σώματα κειμένων, ανάλυσης λόγου (Corpus-assisted Discourse Studies, CADS – Baker, 2006· Gillings κ.ά., 2023· Bednarek κ.ά., 2024).

Η διδακτορική διατριβή της Pearson (1996), που μεταγενέστερα δημοσιεύθηκε και ως μονογραφία (Pearson, 1998), διατυπώνει και οριοθετεί, υπό εφαρμοσμένη οπτική, τη λογική που διέπει τη συγκρότηση και ανάλυση συστηματοποιημένου γλωσσικού υλικού με λεξικογραφική – ορογραφική στόχευση (βλ. Biber & Finegan, 1991· Ooi, 1998· Halliday & Teubert, 2005) και προτείνει ένα μοντέλο συνδυασμού των κλασικών λεξικογραφικών μεθόδων και των δυνατοτήτων που προσφέρουν τα σύγχρονα πληροφοριακά εργαλεία, ιδίως δε τα εργαλεία της ΓΣΚ, για τη λειτουργική σημασιολογική οριοθέτηση των λεξημάτων που αξιοποιούνται ως εστιακά (focal) στη μελέτη των «ειδικών» γλωσσών: μέσω της διεργασίας εντοπισμού του λεξιλογίου, η οποία εστιάζει στις επιτελεστικές πράξεις (illocutionary acts, Austin, 1975), δείχνει το πώς ο ερευνητής μπορεί να συγκεράσει τον σημασιολογικό προσδιορισμό των λέξεων (ή λεξικών κόμβων), αξιολογώντας συνδυαστικά το κειμενικό πλαίσιο στο οποίο αυτές (αυτοί) εντάσσονται και εξετάζονται ως κεντρικές ή ουσιώδεις σημασιακές οντότητες. Η λογική της (κριτικής και συστηματικής) εξέτασης του πλαισίου (context) ανάγεται στο πρόταγμα του Firth (1935, σελ. 40) περί της ανάγκης να εξετάζονται οι «κεντρικές» ή «εστιακές» λέξεις ενός κειμένου, οι οποίες σηματοδοτούν κοινωνικά τη χρήση τους, και είναι καταλυτική στο υπόδειγμα της CADS, εφ' όσον η σχετική διερεύνηση αξιοποιεί εκτεταμένα δείγματα λόγου (και λόγων) των υπό μελέτη γλωσσικών ομάδων – με κοινωνιογλωσσολογική σαφήνεια και διαφάνεια οριοθετημένων. Μεθοδολογικώς συναφείς έρευνες με ορολογική στόχευση (σε ειδικές υπογλώσσες της Ελληνικής) είναι ενδεικτικώς οι: Σαριδάκης (1999), Τζιάφα (2012) και Χριστοπούλου (2024). Στο επίπεδο της ευρετικής μεθοδολογίας, η εργασία αξιοποιεί επίσης τα μοντέλα που αναπτύσσονται στις διδακτορικές διατριβές των Γιούλη (2020), περί της σημασιολογίας του ρήματος στη ΝΕ, και Συμσερίδου (2021), περί της αξιοποίησης της ΓΣΚ στη διδακτική της ειδικής Μετάφρασης.

Στη σύγχρονη εκδοχή τους, η ΓΣΚ και η Κοινωνιογλωσσολογία είναι πεδία ευθέως αλληλένδετα (Baker, 2010). Η προκρινόμενη ευρετική μεθοδολογία καθοδηγείται από τα κειμενικά δεδομένα (corpus-driven – Sinclair, 1991· Kennedy, 1998· Σαριδάκης, 2010· Γούτσος & Φραγκάκη, 2015) και δεν ανατρέχει απλώς σε αυτά (corpus-based) για τον

εντοπισμό και την ερμηνεία των εκάστοτε φαινομένων, χωρίς ωστόσο να οδηγείται σε ακραίες νεο-φιρθιανές προσεγγίσεις που απορρίπτουν καθολικά τη δυνατότητα του ερευνητή να ανατρέχει σε προϋπάρχουσες ταξινομίες ή υποθέσεις για τη γλώσσα (Tognini-Bonelli, 2001· Sinclair, 2004· Teubert & Čermáková, 2007).

Περαιτέρω, σε ακολουθία με τη διεθνή πρακτική, ως μοντέλο διερεύνησης των κοινωνικο-πολιτικών πτυχών της γλωσσικής εκφοράς (των λόγων) της Κοινής Νεοελληνικής (ΚΝΕ) (ενδεικτικά: Hatzidaki & Goutsos, 2017· Hatzidaki & Saridakis, 2020) προτείνεται η «καθοδηγούμενη από τον λόγο ποσοτικοποίηση» (discourse-driven quantification – Goutsos & Hatzidaki, 2017, σελ. 460).

Ειδικότερα σε ό,τι αφορά το υπό μελέτη φαινόμενο και τη δειγματοληψία του υλικού του, επισημαίνονται τα ακόλουθα:

(α) Με βάση τον ορισμό του Ferguson (1959, σελ. 236), η Ελληνική, τουλάχιστον στη διαχρονία της, είναι χαρακτηριστικό παράδειγμα «διγλωσσίας», ήτοι γλωσσικής κατάστασης στην οποία συνυπάρχουν μια «υψηλή» (H) και μια «χαμηλή» (L) ποικιλία. Αν και, ιστορικά, η διάκριση αυτή αναφέρεται αντιστοίχως στην καθαρεύουσα (H) και τη δημοτική (L) (Sifianou, 2003, σελ. 264), ειδικά στο επίπεδο των διαλέκτων και του λεξιλογίου τους εστιάζει στη χρήση συγκεκριμένων τύπων που απαντώνται σε γεωγραφικές ποικιλίες της γλώσσας, με κυρίαρχη τη θεώρηση του Χατζηδάκη ότι αυτές προέρχονται από την Ελληνιστική Κοινή, με την εξαίρεση των Τσακώνικων.

(β) Οι όροι *διάλεκτος* και *ιδίωμα* χρησιμοποιούνται σχεδόν ως συνώνυμοι, με τον πρώτο κατά κανόνα να προσδιορίζει σημαντική διαφοροποίηση μιας ποικιλίας από τη «μέση» γλώσσα, και τον δεύτερο να χρησιμοποιείται για να δείξει διαφορά από τη «μέση» γλώσσα κυρίως στο φωνολογικό επίπεδο, και σε ορισμένες περιπτώσεις στο λεξιλογικό και το μορφοσυντακτικό επίπεδο (Sifianou, 2003, σελ. 267). Ο Χαραλαμπίδης (1992, σελ. 192–193) διαπιστώνει αδυναμία διατύπωσης σαφών κριτηρίων για τη διάκριση μεταξύ κοινών και ιδιωματικών λέξεων, καθώς ακόμη και ο χαρακτηρισμός λέξεων (ή λεξημάτων) ως «συνήθων», «λιγότερο συνήθων», «σπάνιων», κ.λπ. βασίζεται στο υποκειμενικό κριτήριο και το ένστικτο του ερευνητή, ελλείπει (και σήμερα) στατιστικών δεδομένων χρήσης του λεξιλογίου της ΝΕ. Συναφής και η αναφορά της Αγγελοπούλου (2010, σελ. 7) περί δυσδιάκριτου μεταξύ *διαλέκτου* και *ιδιώματος*. Εν τέλει, το ενδιαφέρον βρίσκεται στη στατιστική αποτίμηση και χαρτογράφηση των στοιχείων που συνθέτουν τις γλωσσικές ποικιλίες, με κριτήρια γεωγραφικά (διάλεκτοι, ιδιώματα), κοινωνικά (κοινωνιόλεκτοι) ή λειτουργικά (καταστασιακά ιδιώματα) (Αγγελοπούλου, 2010, σελ. 6).

Κατ' ακολουθία, η λειτουργική αρχή στην οποία βασίστηκε η παρούσα έρευνα για τα *Πολίτικα* είναι «ανοικτή»: το «φαινόμενο» συνθέτουν άξιες μελέτης πραγματώσεις λόγου, είτε ατομικού είτε συλλογικού. Τα *Πολίτικα*, κατά τη διατύπωση της Σελλά-Μάζη (2016), αποτελούν μια «κατάσταση λειτουργικής διγλωσσίας», στην οποία η Ελληνική συνυπάρχει με την Τουρκική (βλ. και Mackridge, 2016, σελ. 2), με αρχαϊσμούς και δάνεια από άλλες γλώσσες, ως αποτέλεσμα αντιστοίχως της επιρροής του Πατριαρχείου (Μήλλας, 1999), της εξωστρέφειας και της κοινωνικής θέσης των Ρωμιών, της επαφής τους με άλλες γλώσσες και κουλτούρες, της πρόσδοσης ειδικού βάρους στη διατήρηση της χρήσης της ποικιλίας ως στοιχείου εθνικής ταυτότητας (Komondouros, 2005· Komondouros & McEntee-Atalianis, 2007), με την ποικιλομορφία των γλωσσικών επιλογών να συσχετίζεται κυρίως με τα πεδία κοινωνικής χρήσης της γλώσσας (Στάθη, 2003, σελ. 456· Τζαβάρας, 2011). Περαιτέρω, τα *Πολίτικα* είναι μια ελάχιστα μελετημένη ποικιλία της ΚΝΕ (Παντελίδης, 2019), παρά την οριοθέτησή της ήδη από τα τέλη του 19ου αιώνα (Hadodo, 2020, σελ. 20, 29, 2025). Στο καθαρά λεξικογραφικό επίπεδο, η μόνη έως σήμερα προσπάθεια (εμπειρική, βιωματική και συνεπώς μη συστηματική σε γλωσσολογικούς όρους) αποθησαύρισης των *Πολίτικων* είναι το λεξικό του Ζαχαριάδη (2014).

1 Σώμα κειμένων, μεθοδολογία ανάλυσης

Για τις ανάγκες της έρευνας, συγκροτήθηκε αντιπροσωπευτικό της ποικιλίας και του κειμενικού είδους¹ σώμα τριών λογοτεχνικών κειμένων, από κωνσταντινουπολίτισσες συγγραφείς, που έχουν ως κεντρική θεματολογία την ιστορική, κοινωνική και πολιτισμική πραγματικότητα της ελληνορθόδοξης κοινότητας της Κωνσταντινούπολης, της περιόδου 1950-1964.

¹ Επιγραμματικά, η αντιπροσωπευτικότητα του δείγματος που επιλέχθηκε για τη συγκρότηση του ΣΚ της έρευνας εδράζεται σε εξωτερικά και εσωτερικά κριτήρια (Sinclair & Ball, 1996· McEnery & Wilson, 2001, σελ. 31): τα μεν πρώτα συσχετίζουν το βασικό κοινωνιογλωσσικό προφίλ των συγγραφέων, το κειμενικό είδος, τη θεματική αναφορά και τη χρονικότητά της, όπως και τη δηλωτική έκθεση της γλωσσικής επιτέλεσης (Austin, 1975, p. 161· Pearson, 1998, p. 116· Saridakis, 2013, p. 199-200), όπως αυτή αποτυπώνεται στους προλόγους των λογοτεχνικών έργων (σχετική και η παρατήρηση του Hadodo (2020, σελ. 1) περί των μηχανισμών δόμησης της ταυτότητας των Ελλήνων της Πόλης), τα δε δεύτερα σταθμίστηκαν ενδογλωσσικά, με πιλοτική στατιστική ανάλυση του λεξικογραφικού προφίλ των κειμενικών δειγμάτων.

Πίνακας 1. Το ΣΚ της έρευνας.

Τίτλος βιβλίου	Συγγραφέας	Έτος	Έκδοση	Έκταση
[01] Οι ποδηλάτισσες	Ρ. Σταθοπούλου	2022	Τσουκάτος	90065
[02] Ο βασιλιάς των αριθμών	Ρ. Σταθοπούλου	2006	Ωκεανίδα	93870
[03] Άντε πάλι... ξεριζώνεται ο ελληνισμός από την Πόλη (1958-1968)	Α. Κωνσταντινίδου-Κλουκίνα	2011	Ίκαρος	174756
Σύνολο				358691 λέξεις (16008 φράσεις)

Το ΣΚ αναλύθηκε ως προς τα λεξιλογικά (και κατ' επέκταση λεξικογραμματικά) του στοιχεία (αρχικά λήμματα: 20.557). Εντοπίστηκαν και λημματοποιήθηκαν τα δάνεια ουσιαστικά (N=372). Η στάθμιση του φαινομένου του δανεισμού έγινε με βάση τη στατιστική σημαντικότητα της εμφάνισης στο ΣΚ των συγκεκριμένων λεξημάτων, σε σύγκριση με ΣΚ γενικής γλώσσας (EITenTen). Περαιτέρω, το φαινόμενο του δανεισμού σταθμίστηκε (και δεικτοδοτήθηκε καταλλήλως) με αντιβολή στα κύρια σύγχρονα γενικά λεξικά της ΝΕ (Μπαμπινιώτη, Ιδρύματος Τριανταφυλλίδη και Ακαδημίας Αθηνών) και στο λεξικό του κωνσταντινουπολίτικου ιδιώματος (Ζαχαριάδης, 2014). Τα δάνεια ουσιαστικά επισημειώθηκαν ως προς την κοινωνιοσημειωτική τους υπόσταση (Halliday, 1978, 1997· Cobley & Randviir, 2009), ειδικότερα δε ως προς την πτυχή του κώδικα² με κριτική³

² Κατά τον Bernstein (2003, σελ. 58–62): η επιλογή της γλώσσας είναι ουσιώδης και αποτελεί ενδείκτη της κοινωνικής ταυτότητας, δεδομένου ότι χρησιμοποιείται για τη δημιουργία κοινωνικών δομών, σε μια γενική θεωρία πολιτισμικής ανταλλαγής (Moore, 2013, σελ. 24) και κατ' ακολουθία αντανάκλα και συγκεκριμενοποιεί τις υποθέσεις περί διακριτών κοινωνικών ομάδων, ενώ οι σχέσεις που αναπτύσσονται στο εσωτερικό μίας εκάστης ομάδας «επηρεάζουν τον τρόπο με τον οποίο η ομάδα αυτή χρησιμοποιεί τη γλώσσα και τον τύπο λόγου που αναπτύσσει» (Clark, 2013, σελ. 113). Ο Bernstein διακρίνει μεταξύ «περιορισμένου» και «επεξεργασμένου» κώδικα. Ο πρώτος εντοπίζεται σε μη τυπικές και καθημερινές επικοινωνιακές περιστάσεις, συσχετιζόμενος στενά με το καταστασιακό πλαίσιο, και χαρακτηρίζεται από απουσία υφολογικής ποικιλομορφίας και από έμφαση στη συμμετοχή του ομιλητή στην (εν λόγω) κοινωνική ομάδα, από απλές συντακτικές δομές και από συχνή χρήση, στην προφορική επικοινωνία, εξωλεκτικών στοιχείων και επιβεβαιωτικών ερωτήσεων στο τέλος μίας σηματοδοτικής ενότητας (π.χ. «έτσι δεν είναι;»). Ο επεξεργασμένος κώδικας χρησιμοποιείται σε συμβατικοποιημένες περιστάσεις επικοινωνίας, εξαρτάται λιγότερο από το πλαίσιο, ανατρέχει σε μεγάλο εύρος υφολογικών διαφοροποιήσεων, και χαρακτηρίζεται από συχνή χρήση της παθητικής φωνής, αυξημένη αναλογία επιθέτων και προσδιοριστικών και ενίοτε σύνθετες συντακτικές δομές. Αν και η συσχέτιση των δύο αυτών τύπων κώδικα με την «κοινωνική τάξη» αμφισβητείται ως προς την καθολικότητά της, ως «μη άμεση, δεδομένης της παρεμβολής και άλλων παραγόντων όπως το πλαίσιο στο οποίο μαθαίνεται η γλώσσα και ο τρόπος με τον οποίο

συνεξέταση του κειμενικού, συμφραστικού περιβάλλοντος (co-text, concordances – Partington, 2004) των κύριων ονοματικών (δανείων) λημμάτων, ακολουθώντας το υπόδειγμα της καθοδηγούμενης από τον λόγο ποσοτικοποίησης (Goutsos & Hatzidaki, 2017)⁴.

Οι κατηγορίες που αναπτύσσονται, με ποιοτική συστηματοποίηση των ευρημάτων (πβ. Ide & Rutejovsky, 2017), ακολουθώντας το γενικό κοινωνιογλωσσολογικό υπόδειγμα της Ανάλυσης Λόγου (ενδ. Saridakis & Mouka, 2018) είναι: (α) Κοινωνική υπόσταση/θέση της κατονομαζόμενης οντότητας (προσώπου, αντικειμένου, κατάστασης - SOCST), (β) Καθημερινή ζωή στο πλαίσιο της κοινότητας (DAYL, με πέντε διακριτές εστιακές υποκατηγορίες), (γ) Προσωπικά χαρακτηριστικά (λεξήματα με κυρίως αξιολογικό περιεχόμενο και με αναφορά σε πρόσωπα στο πλαίσιο της κοινότητας - PERSTRAIT), (δ) Επάγγελμα (PROF), (ε) Λοιπά λεξήματα (μη αυτοτελώς κατηγοριοποιούμενα εκτός κειμενικού περιβάλλοντος - OTHER).

Αντιπαραβάλλοντας τη σύνθεση της κοινωνιοσημειωτικής χρήσης, και με σκοπό μια πρώτη χαρτογράφηση της κοινωνιογλωσσολογικής πραγματικότητας της Κωνσταντινούπολης και των βασικών κοινωνικοπολιτικών συντελεστών διαμόρφωσης των ιδιαιτεροτήτων της

δομείται η οικογενειακή ζωή» (Crystal, 1997, σελ. 40), χρησιμοποιείται συχνά για τη χαρτογράφηση της κοινωνικής προέλευσης των ομιλητών (Chandler, 2007, σελ. 153–154).

³ Η πτυχή της (κριτικής) προσέγγισης και ερμηνείας του (μακρο) κειμενικού περιβάλλοντος, ως ποιοτικής μεταβλητής της διεργασίας εκκινεί από την παραδοχή του πεδίου της Κριτικής Ανάλυσης Λόγου (CDA, ΚΑΛ) ότι η γλώσσα, πέρα από σημειωτικό σύστημα, είναι μορφή κοινωνικής πρακτικής που αναπαριστά και εν τέλει καθορίζει την κοινωνική πραγματικότητα, η δε κειμενική ανάλυσή της εκκινεί από την περιγραφή του τρόπου έκφρασης των κοινωνικών ζητημάτων στον λόγο για να οδηγηθεί στην ερμηνεία τους (κοινωνική ανάλυση) (Μπουκάλα & Στάμου, 2020, σελ. 11). Κατ' ακολουθία, οι γλωσσικές πράξεις, και ευρύτερα οι λόγοι των ατόμων και των κοινωνικών ομάδων, παράγουν και συντηρούν σχέσεις ισχύος, κυριαρχίας, διάκρισης και διαφοροποίησης (van Dijk, 1977· Fairclough, 1992· Wodak & Meyer, 2001). Κοινός παρονομαστής των προσεγγίσεων στο πεδίο της ΚΑΛ είναι, εν ολίγοις, η έμφαση στις σχέσεις ισχύος και στις αλλαγές που παρατηρούνται στην κοινωνική σφαίρα στο πολιτικό, κοινωνικό και πολιτισμικό επίπεδο, ώστε, ερμηνευτικά ο ερευνητής να οδηγηθεί στην «ανάγνωση» των ιδεολογιών και της ισχύος (Wodak, 2024, σελ. 31–32).

⁴ Στην πράξη, ο ερευνητής δεν στοχεύει στην επαλήθευση ή διάψευση μιας προϋπάρχουσας ερμηνευτικής ταξινόμησης, αλλά στην αξιοποίησή της, στο μέτρο του δυνατού, ως εργαλείου με ανοικτή ερμηνευτική προοπτική: «Η CADS συγκεντρώνει παρεμφερείς οπτικές. Στοιχείο-κλειδί είναι η γλωσσική δομή. Οι οπτικές αυτές βρίσκονται στο επίκεντρο της ΓΣΚ και της Ανάλυσης Λόγου. Η πρώτη εντοπίζει τις κανονικότητες, με το αξιολογικό φορτίο των λεξιλογικών συσχετίσεων, και η δεύτερη ερμηνεύει τον τρόπο με τον οποίο συστηματοποιούνται αυτές οι συσχετίσεις στο κοινωνικοπολιτικό πλαίσιο. Συνεργιστικά, μπορούν να δώσουν μια καθαρή εικόνα του τρόπου με τον οποίο συστηματοποιούνται οι λόγοι και καθίστανται κυρίαρχοι, μέσω της επαναλαμβανόμενης και αυξανόμενης χρήσης» (Stubbs, 2002, σελ. 215 με αναφορές στα: Baker, 2006, σελ. 13· Gillings κ.ά., 2023, σελ. 6).

συγκεκριμένης γλωσσικής ποικιλίας, η μελέτη τριγωνίζεται μεθοδολογικά, συνεξετάζοντας ευρήματα από στοχευμένες συνεντεύξεις με τρεις φυσικές ομιλήτριες της ποικιλίας – μία εκ των δύο συγγραφέων των λογοτεχνικών έργων του ΣΚ (Κωνσταντινίδου-Κλουκίνα), και δύο διδάσκουσες τουρκικής γλώσσας τού ΕΚΠΑ.

Η διαδικασία της ανάλυσης και κριτικής ερμηνείας των λεξιλογικών ευρημάτων συνοψίζεται στο ακόλουθο σχήμα:

[ΠΟΣΟΤΙΚΗ ΑΝΑΛΥΣΗ : Λέξεις-κλειδιά →λημματοποίηση, γραμματική
επισημείωση→κοινωνιοσημειωτική επισημείωση →στατιστική επεξεργασία →ανάλυση συμφράσεων
(KWIC)
→λημματολόγιο →ερμηνεία στατιστικών ευρημάτων].
→[ΠΟΙΟΤΙΚΗ ΑΝΑΛΥΣΗ : ΣΥΝΝΕΤΕΥΞΕΙΣ (σύνθεση και αποτελέσματα)]→

Διάγραμμα 1. Ροή ανάλυσης και κριτικής ερμηνείας λεξιλογικών ευρημάτων.

2 Τεχνική επεξεργασία

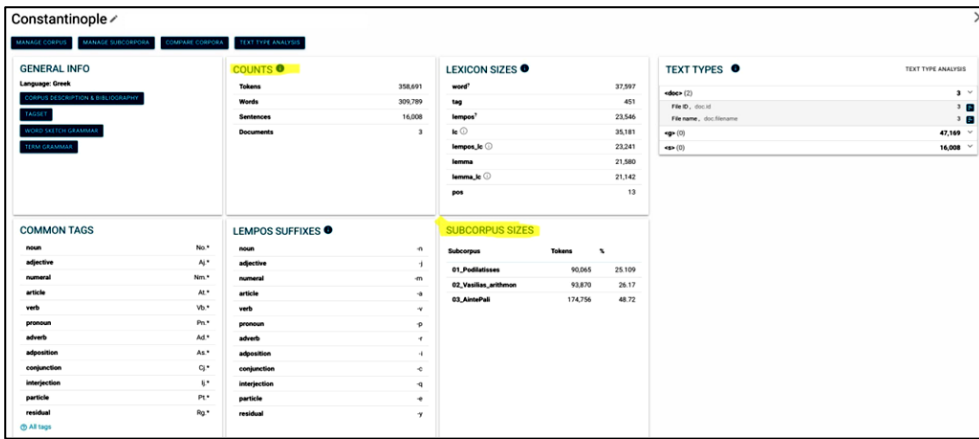
Διαγραμματικά, η διεργασία τεχνικής επεξεργασίας και συστηματοποίησης του πρωταρχικού γλωσσικού υλικού και σύνθεσης του ορολογικού πόρου είναι εν συνόψει η ακόλουθη:

Ψηφιοποίηση και κανονικοποίηση →συγκρότηση και ανάλυση ΣΚ →στατιστική επεξεργασία,
οπτικοποίηση →μεταγραφές IPA → σύνθεση γλωσσαρίου →βάση δεδομένων

Διάγραμμα 2. Ροή επεξεργασίας και συστηματοποίησης υλικού.

Τα κείμενα, μετά από σάρωση, ψηφιοποιούνται σε αναγνώσιμη μορφή χρησιμοποιώντας το λογισμικό Tesseract OCR της Google⁵, μετατρέπονται σε μονοτονική γραφή με χρήση script σε Linux ('sed' με πίνακα αντικατάστασης) και διορθώνονται ορθογραφικά σε επεξεργαστή κειμένου. Το ΣΚ συγκροτείται και αναλύεται στο SketchEngine.

⁵ Διαθέσιμο στο github <<https://bit.ly/49Jav2X>>.



Εικόνα 1. Η δομή του ΣΚ στο SketchEngine.

Η στάθμιση και οπτικοποίηση των βασικών στατιστικών ευρημάτων γίνεται στο LibreOffice Calc. Τα βασικά στατιστικά δεδομένα συλλέγονται σε βιβλίο εργασίας, με 34 διακριτά φύλλα.

Type	RF_S03	Score_S03	F_SC	Integer_check	RF_SC	Score_total	F_RFC	RF_RC	DOCF_S03(a)	DOCF_RC(N)	ARF_SC	ARF_RC	F_SC_H0	Integer_check	F_RC_H0	Fm_S03_H0
15400_αυτοία	22,89	268,87	4	4	11,15	102,23	26	0,01	1	0,00	2,24	15,39	4	4	26	11,15
15760_σπυρίδι	17,17	162,14	3	3	8,36	79,50	18	0,01	1	0,00	1,43	8,51	3	3	18	8,36
15760_σπυρίδι	11,44	111,83	2	2	5,58	54,98	9	0,00	1	0,00	1,01	4,86	2	2	9	5,58
15848_σπυρίδι	11,44	112,81	2	2	5,58	55,37	7	0,00	1	0,00	1,77	3,41	2	2	7	5,58
16126_σπυρίδι	11,44	115,45	2	2	5,58	56,76	6	0,00	1	0,00	1,54	0,00	2	2	6	5,58
16441_σπυρίδι	28,61	285,56	5	5	13,94	139,39	2	0,00	1	0,00	1,61	1,00	5	5	2	13,94
16531_σπυρίδι	17,17	166,10	3	3	8,36	81,42	11	0,00	1	0,00	1,38	7,56	3	3	11	8,36
17866_σπυρίδι	5,72	56,79	2	2	5,58	55,37	7	0,00	2	0,00	1,78	4,70	2	2	7	5,58
17909_σπυρίδι	28,61	286,59	5	5	13,94	139,89	1	0,00	1	0,00	2,68	1,00	5	5	1	13,94
17949_σπυρίδι	28,61	225,51	6	6	13,94	110,27	76	0,03	1	0,00	1,61	40,72	5	5	76	13,94
17949_σπυρίδι	11,44	113,57	2	2	5,58	55,56	4	0,00	1	0,00	1,58	2,87	2	2	4	5,58
17949_σπυρίδι	57,22	561,13	10	10	27,88	273,89	6	0,00	1	0,00	4,03	2,21	10	10	6	27,88
17866_σπυρίδι	17,17	170,22	3	3	8,36	83,44	4	0,00	1	0,00	1,34	2,84	3	3	4	8,36
17883_σπυρίδι	11,44	115,03	2	2	5,58	56,56	1	0,00	1	0,00	1,32	1,00	10	10	44	27,88
17883_σπυρίδι	45,76	397,37	8	8	22,30	194,04	42	0,02	1	0,00	3,69	34,07	8	8	42	22,30
18028_σπυρίδι	5,72	56,22	2	2	5,58	56,76	0	0,00	2	0,00	1,38	0,00	2	2	0	5,58
18089_σπυρίδι	11,44	114,21	2	2	5,58	56,15	3	0,00	1	0,00	1,11	2,23	2	2	3	5,58
18197_σπυρίδι	11,44	109,68	2	2	5,58	58,16	120	0,06	1	0,00	1,37	79,61	2	2	120	5,58
18349_σπυρίδι	0,00	0,00	2	2	5,58	54,79	10	0,00	2	0,00	1,68	7,13	2	2	10	5,58
18411_σπυρίδι	5,72	51,07	2	2	5,58	49,78	39	0,01	2	0,00	1,84	23,61	2	2	39	5,58
18438_σπυρίδι	0,00	0,00	2	2	5,58	23,81	365	0,14	2	0,00	1,13	209,66	2	2	365	5,58
18438_σπυρίδι	5,72	55,24	4	4	11,15	106,76	15	0,01	2	0,00	2,00	7,86	4	4	15	11,15
18442_σπυρίδι	17,17	170,83	3	3	8,36	83,74	3	0,00	1	0,00	1,79	2,41	4	4	11	11,14
18442_σπυρίδι	11,44	115,45	2	2	5,58	56,76	10	0,00	1	0,00	1,47	0,00	2	2	10	5,58
18446_σπυρίδι	11,44	109,69	2	2	13,94	130,11	22	0,01	2	0,00	2,82	19,44	6	6	22	16,71
19226_σπυρίδι	17,17	37,99	3	3	8,36	18,18	1017	0,37	1	0,01	1,57	514,84	3	3	1017	8,36

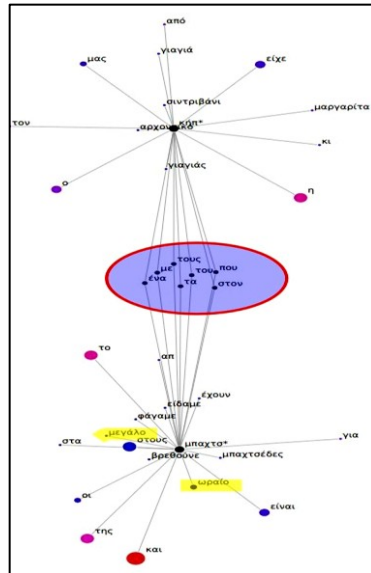
Εικόνα 2. Κεντρικός πίνακας υπολογισμού στατιστικών δεδομένων (Calc).

Σε διακριτό πίνακα εξάγονται το σύνολο των συμφράσεων (concordances) των εστιακών λεξημάτων, σε διασύνδεση με το βασικό βιβλίο εργασίας.

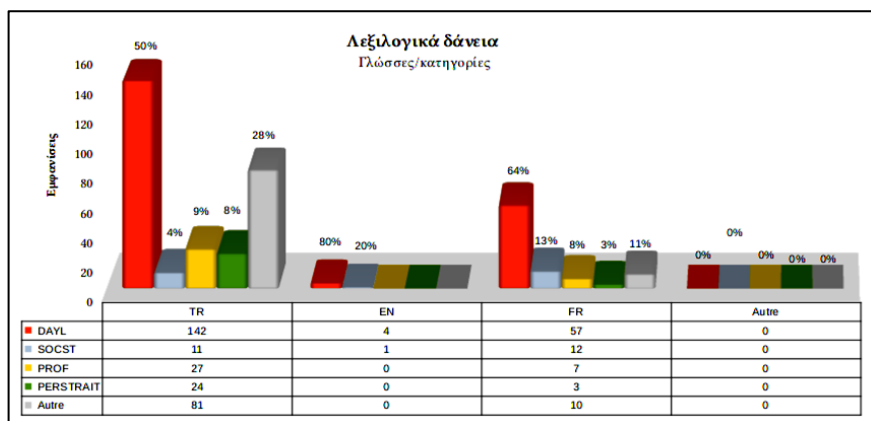


Εικόνα 3. Παράδειγμα συμφραστικού πίνακα κομβικής (εστιακής) λέξης.

Για την ανάλυση των λεξικών παραθεμάτων και της «γεωμετρίας» τους (πέραν των βασικών στατιστικών που εξάγονται από το SketchEngine, χρησιμοποιείται το εργαλείο LancsBox του Πανεπιστημίου του Lancaster (Brezina & Platt, 2023).



Εικόνα 4. Λεξιλογικό-σημασιολογικό δίκτυο κοινών παραθεμάτων (c-collocates) – LancsBox.



Εικόνα 5. Ονοματικά δάνεια/κοινωνιοσημειωτικές κατηγορίες.

Τα εστιακά λεξήματα μεταγράφονται στο διεθνές φωνολογικό αλφάβητο (IPA/APA): χρησιμοποιούμε Python script (864 γραμμές κώδικα) το οποίο αναπτύξαμε ad hoc, με υποστήριξη στην αποσφαλμάτωση του κώδικα από το ChatGPT⁶.

⁶ Το script εκπαιδεύτηκε στο ελληνικό ορθογραφικό λεξικό (GNU: <<https://packages.debian.org/sid/aspell>>), και σε αυτό ενσωματώθηκαν βιβλιοθήκες ανάγνωσης διαφόρων τύπων αρχείων (txt, csv, odt, docx, ods, xlsx, pdf, κ.λπ.), και, τέλος, προσαρμόστηκε στο γραφικό περιβάλλον Gtk, ώστε να λειτουργεί ως αυτοδύναμη (stand-alone) εφαρμογή σε περιβάλλον Linux.

```
import gi
import re
import csv
from odf.opendocument import load
from odf.text import P, H
from docx import Document
import texttract
from pdfminer.high_level import extract_text
from pyexcel_ods import get_data

gi.require_version('Gtk', '3.0')
from gi.repository import Gtk, Gdk

# Mapping from Greek characters, diphthongs, and accented characters to IPA
greek_to_ipa = {
    'αί': 'aj', 'αϊ': 'aj', 'αί': 'aj', 'εϊ': 'ej', 'εϊ': 'ej',
    'οϊ': 'oj', 'οϊ': 'oj', 'υϊ': 'yj', 'υϊ': 'yj',
    'εύ': 'eu', 'ευ': 'eu',
    'ού': 'u', 'ου': 'u', 'αί': 'e', 'εϊ': 'i', 'εϊ': 'ei', 'οι': 'i', 'υι': 'i', 'αυ': 'au', 'ευ': 'eu',
    'μπ': 'b', 'ντ': 'd', 'γκ': 'g', 'τζ': 'dz', 'το': 'ts',
    'α': 'a', 'β': 'v', 'γ': 'x', 'δ': 'ð', 'ε': 'e',
    'ζ': 'z', 'η': 'i', 'θ': 'θ', 'ι': 'i', 'ξ': 'i',
    'κ': 'k', 'λ': 'l', 'μ': 'm', 'ν': 'n', 'ξ': 'ks', 'ο': 'o',
    'π': 'p', 'ρ': 'r', 'σ': 's', 'ς': 's', 'τ': 't', 'υ': 'i',
    'ϕ': 'i', 'φ': 'f', 'χ': 'x', 'ψ': 'ps', 'ω': 'o',
    'ά': 'a', 'έ': 'e', 'ή': 'i', 'ί': 'i', 'ό': 'o', 'ύ': 'i', 'ώ': 'o',
    'ϊ': 'i', 'ϋ': 'i'
}

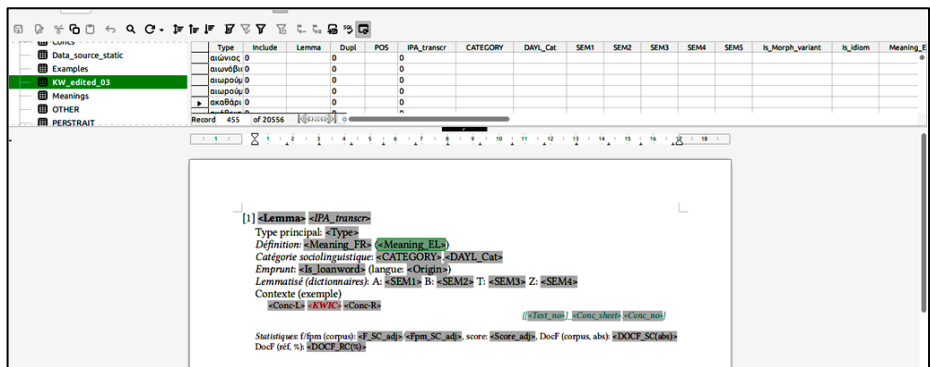
# Common Greek abbreviations
abbreviations = [
    'κ.', 'π.', 'χ.', 'β.', 'δ.', 'μ.', 'λ.', 'εξ.', 'κα.', 'ελ.', 'πχ.', 'κτλ.', 'κλπ.'
]

# Characters with accents
toned_chars = 'άέήιούώ'

def replace_abbreviations(text):
    """
    Replaces abbreviations in the text.
    """
    for abbrev in abbreviations:
        text = re.sub(fr'\b{re.escape(abbrev)}\b', abbrev.replace('.', ''), text)
    return text
```

Εικόνα 6. Τμήμα κώδικα μεταγραφής σε IPA/APA (python).

Τέλος, το επισημειωμένο υλικό συγκεντρώνεται σε λεξικογραφικό πόρο (δομή τελικού πίνακα δεδομένων: MySQL, XML), υπό μορφή αυτόματα δημιουργούμενου (βάσει προτύπου) λεμματολογίου που περιλαμβάνει παραδείγματα χρήσης και βασικά στατιστικά δεδομένα χρήσης.



Εικόνα 7. Πρότυπο αυτοματοποίησης δημιουργίας λεμματολογίου (Libre Office Writer).

[11] *απτάλης* /a'ptalis/

Type principal: *απτάλης*

Définition: imbecile (βλάκας)

Catégorie sociolinguistique: PERSTRAIT.

Langue d'origine: TR

Lemmatisé (dictionnaires): A: 1 B: 0 T: 1 Z: 0

Contexte (exemple)

Ο Αρτέμης έμεινε να τον κοιτάζει για λίγα δευτερόλεπτα σαν *απτάλης* χωρίς να καταλαβαίνει τι εννοούσε ο φίλος του.

[[3]_11.6]

Εικόνα 8. Παράδειγμα βασικής λημματογράφησης.

3 Κύρια ευρήματα, ερμηνευτική σύνθεση

Τα *Πολίτικα* χρησιμοποιούνται ως ισχυρός μηχανισμός ταυτοτικής κατασκευής, καλύπτοντας την ανάγκη των μελών της κοινότητας να αυτοχαρακτηρισθεί ως «ειδική ομάδα», με τα δικά της πολιτισμικά χαρακτηριστικά που τη διακρίνουν από τους Τούρκους φυσικούς ομιλητές με τους οποίους η κοινότητα συμβιώνει, γεωγραφικά, επαγγελματικά και κοινωνικά.

Η χρήση των ονοματικών δανείων είναι στρωματοποιημένη: η διαφορά στην προέλευση των δανείων παραπέμπει, με βάση την ανάλυση του παραθετικού προφίλ και των συμφράσεων των εστιακών λημμάτων και κατά συνέπεια με βάση τη σχηματοποίηση στο επίπεδο της σημασιολογικής προτίμησης και προσωδίας, σε διάκριση μεταξύ «υψηλής» και «χαμηλής» ποικιλίας (Ferguson, 1959· Fishman, 1967· Σελλά-Μάζη, 2016).

Συγκεκριαστικά, η εικόνα που σχηματίζεται ως προς την κοινωνική σημείωση στο πλαίσιο της καταγεγραμμένης εκφοράς αντιστοιχεί σε κατάσταση λειτουργικής διγλωσσίας (Σελλά-Μάζη, 2016): σε αυτήν, τα λεξήματα, σε όλες τις γραμματικές κατηγορίες, που προέρχονται από την Τουρκική (303 δάνεια, 65,02%) αντιστοιχούν στη «χαμηλή» ποικιλία και χρησιμοποιούνται για τον αρνητικό χαρακτηρισμό των ονοματιζόμενων προσώπων, οντοτήτων και καταστάσεων. Αντιστρόφως, η χρήση δανείων από δυτικές γλώσσες (FR [93,19,95%], IT [65,13,94%], EN [5,1,07%]) αντικατοπτρίζει υψηλό μορφωτικό, κοινωνικό και επαγγελματικό επίπεδο, όπως και θετική αξιολόγηση (White, 2015) των κατονομασιών (Wodak, 2015).

Βιβλιογραφικές αναφορές

- Αγγελοπούλου, Δ. (2010). *Το γορτυνιακό ιδίωμα και η χρήση του στην εκπαίδευση* [Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών (ΕΚΠΑ)]. <https://doi.org/10.12681/eadd/25395>
- Γιούλη, Π. (2020). *Το σημασιολογικό πεδίο των συναισθημάτων: Ταξινόμηση των ρημάτων της Νέας Ελληνικής που δηλώνουν συναίσθημα* [Διδακτορική διατριβή, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών (ΕΚΠΑ), Σχολή Οικονομικών και Πολιτικών Επιστημών, Τμήμα Τουρκικών Σπουδών και Σύγχρονων Ασιατικών Σπουδών]. <https://doi.org/10.12681/eadd/50114>
- Γούτσος, Δ., & Φραγκάκη, Γ. (2015). *Εισαγωγή στη γλωσσολογία σωμάτων κειμένων*. Κάλλιπος. <https://repository.kallipos.gr/handle/11419/1932>
- Ζαχαριάδης, Ν. (2014). *Λεξικό του κωνσταντινοπολίτικου γλωσσικού ιδιώματος*. Γαβριηλίδης.

- Μήλλας, Η. (1999). *Το κωνσταντινουπολίτικο ιδίωμα*. <https://www.herkulmillas.com/el/hm-makaleleri/glossa-logotexnia/643-to-konstantinoupolitiko-idioma.html>
- Μπουκάλα, Σ., & Στάμου, Α. (2020). Εισαγωγικό σημείωμα. Στο Σ. Μπουκάλα & Α. Στάμου (Επιμ.), *Κριτική Ανάλυση Λόγου. (Απο)δομώντας την ελληνική πραγματικότητα* (σελ. 11–17). Νήσος.
- Παντελίδης, Ν. (2019). Σελίδες από την ιστορία της ελληνικής γλώσσας: Η περίπτωση της Κωνσταντινούπολης. Στο Α. Αρχάκης (Επιμ.), *Γλωσσική ποικιλία. Μελέτες αφιερωμένες στην Αγγελική Ράλλη* (σελ. 453–480). Κάπα εκδοτική.
- Σαριδάκης, Ι. (1999). *Μεθοδολογική προσέγγιση της ορογραφικής διαδικασίας. Η δόμηση αγγλοελληνικού καταλόγου αμυντικών όρων* [Διδακτορική διατριβή, Ιόνιο Πανεπιστήμιο, Τμήμα Ξένων Γλωσσών, Μετάφρασης και Διερμηνείας]. <https://doi.org/10.12681/eadd/15688>
- Σαριδάκης, Ι. Ε. (2010). *Σώματα Κειμένων και Μετάφραση. Θεωρία και Εφαρμογές*. Παπαζήσης.
- Σελλά-Μάζη, Ε. (2016). *Διγλωσσία, εθνική ταυτότητα και μειονοτικές γλώσσες*. Λειμών.
- Στάθη, Π. (2003). Θέσεις και αντιθέσεις σχετικά με το κωνσταντινουπολίτικο ιδίωμα. *Πρακτικά του τέταρτου διεθνούς συνεδρίου Νεοελληνικής Διαλεκτολογίας (2001)*, 451–460.
- Συμσερίδου, Ε. (2021). *Η χρήση των ηλεκτρονικών σωμάτων κειμένων στη διδακτική της ειδικής μετάφρασης: Μία θεωρητική και πρακτική προσέγγιση* [Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης (ΑΠΘ), Σχολή Φιλοσοφική, Τμήμα Γαλλικής Γλώσσας και Φιλολογίας, Τομέας Μετάφρασης και Μεταφρασεολογίας]. <https://doi.org/10.12681/eadd/49689>
- Τζαβάρας, Ξ. (2011). *Λεξικό του ιμβριακού ιδιώματος. Ερμηνευτικό-ετυμολογικό*. Σύλλογος Ιμβριών Αθηνών.
- Τζιάφα, Ε. (2012). *Μελέτη της ειδικής γλώσσας του χρηματιστηρίου με βάση σώματα κειμένων και στόχο την αυτόματη μετάφραση* [Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης (ΑΠΘ), Σχολή Φιλοσοφική, Τμήμα Αγγλικής Γλώσσας και Φιλολογίας]. <https://doi.org/10.12681/eadd/35252>
- Χαραλαμπίδης, Χ. (1992). *Νεοελληνικός Λόγος. Μελέτες για την Γλώσσα, τη Λογοτεχνία και το Ύψος*. Νεφέλη.
- Χριστοπούλου, Α. (2024). *Ειδική γλώσσα και ορολογία της δημόσιας διοίκησης: Αντιπαραβολική μελέτη ελληνικής-γαλλικής με χρήση σωμάτων κειμένων στους τομείς της απασχόλησης και της οικονομίας* [ΕΚΠΑ]. <https://doi.org/10.12681/eadd/57022>
- Austin, J. L. (1975). *How to do things with words*. Clarendon Press.
- Baker, P. (2006). *Using Corpora in Discourse Analysis*. Continuum.
- Baker, P. (2010). *Sociolinguistics and Corpus Linguistics*. Edinburgh University Press.
- Bednarek, M., Schweinberger, M., & Lee, K. K. H. (2024). Corpus-based discourse analysis: From meta-reflection to accountability. *Corpus Linguistics and Linguistic Theory*, 20(3), 539–566. <https://doi.org/10.1515/cilt-2023-0104>
- Bernstein, B. (2003). *Theoretical Studies Towards a Sociology of Language* (2nd edition). Routledge.
- Biber, D., & Finegan, E. (1991). On the Exploitation of Computerized Corpora in Variation Studies. Στο K. Aijmer & B. Altenberg (Επιμ.), *English Corpus Linguistics* (σελ. 204–220). Routledge.
- Brezina, V., & Platt, W. (2023). *LancsBox* [Software]. <http://lancsbox.lancs.ac.uk>
- Chandler, D. (2007). *Semiotics. The Basics* (2nd edition). Routledge.
- Clark, U. (2013). *Language and Identity in Englishes*. Routledge. <https://doi.org/10.4324/9780203552537>
- Cobley, P., & Randviir, A. (2009). Introduction: What is sociosemiotics? *Semiotica*, 2009(173), 1–39. <https://doi.org/10.1515/SEMI.2009.001>
- Crystal, D. (1997). *The Cambridge encyclopedia of language* (2nd ed). Cambridge University Press.
- Dijk, T.A. van. (1977). *Text and Context. Explorations in the Semantics and Pragmatics of Discourse*. Longman.
- Fairclough, N. (1992). Linguistic and intertextual analysis within discourse analysis. *Discourse & Society*, 3(2), 193–217. <https://doi.org/10.1177/0957926592003002004>
- Ferguson, C. (1959). Diglossia. *Word*, 15, 325–340.

- Firth, J. R. (1935). The Technique of Semantics. *Transactions of the Philological Society*, 34(1), 36–73. <https://doi.org/10.1111/j.1467-968X.1935.tb01254.x>
- Fishman, J. (1967). Bilingualism with and without diglossia, diglossia with and without bilingualism. *Journal of Social Issues*, XXIII(2), 29–37. <https://doi.org/10.1111/j.1540-4560.1967.tb00573.x>
- Gillings, M., Mautner, G., & Baker, P. (2023). *Corpus-Assisted Discourse Studies*. CUP.
- Goutsos, D., & Hatzidaki, O. (2017). Making sense of the Greek crisis. Στο O. Hatzidaki & D. Goutsos (Επιμ.), *Greece in Crisis. Combining critical discourse and corpus linguistics perspectives* (τ. 70, σελ. 457–466). John Benjamins. <https://doi.org/10.1075/dapsac.70.13gou>
- Hadodo, M. J. (2020). *Cosmopolitan Constantinopolitans* [PhD thesis]. University of Pittsburgh.
- Hadodo, M. J. (2025). From “Turkish Seeds” to “Greekier than the Greeks”. Στο M. J. Hadodo, P. Karatsareas, & E. Ioannidou (Επιμ.), *Greek in Minoritized Contexts* (1η έκδ., σελ. 25–48). Routledge. <https://doi.org/10.4324/9781003269694-3>
- Halliday, M. (1978). *Language as social semiotic. The social interpretation of language and meaning*. Arnold.
- Halliday, M. (1997). Language in a social perspective. Στο N. Coupland & A. Jaworski (Επιμ.), *Sociolinguistics: A reader*, σελ. 31–38. Palgrave Macmillan.
- Halliday, M., & Teubert, W. (Επιμ.). (2005). *Lexicology and corpus linguistics: An introduction* (Reprint). Continuum Books.
- Hatzidaki, O., & Goutsos, D. (2017). The discourses of the Greek crisis. Στο O. Hatzidaki & D. Goutsos (Επιμ.), *Greece in crisis. Combining critical discourse and corpus linguistics perspectives* (σελ. 3–41). John Benjamins.
- Hatzidaki, O., & Saridakis, I. E. (2020). Discourses of aggression in Greek digitally-mediated communication. An overview of published research. Introduction. *Journal of Language Aggression and Conflict*, 8(2), 147–155. <https://doi.org/10.1075/jlac.00043.edi>
- Ide, N., & Pustejovsky, J. (Επιμ.). (2017). *Handbook of linguistic annotation*. Springer.
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. Longman.
- Komondouros, M. (2005). *Language Attitudes and use in the Rum Community of Istanbul* [MA thesis in Applied Linguistics]. University of Reading.
- Komondouros, M., & McEntee-Atalianis, L. (2007). Language Attitudes, Shift and the Ethnolinguistic Vitality of the Greek Orthodox Community in Istanbul. *Journal of Multilingual and Multicultural Development*, 28(5), 365–384. <https://doi.org/10.2167/jmmd483.1>
- Mackridge, P. (2016). Το πολιτικό λεξιλόγιο: Από τους Φαναριώτες μέχρι τη σημερινή εποχή. Παρουσίαση του *Λεξικού του κωνσταντινουπολίτικου γλωσσικού ιδιώματος* του Νίκου Ζαχαριάδη στην Ημερίδα «Η Ελληνική της Κωνσταντινούπολης: παρελθόν, παρόν και μέλλον», Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Τουρκικών Σπουδών και Σύγχρονων Ασιατικών Σπουδών, Αθήνα, 18 Μαΐου 2016.
- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics*. Edinburgh University Press.
- Moore, R. (2013). *Basil Bernstein: The thinker and the field*. Routledge. <https://doi.org/10.4324/9780203818251>
- Ooi, V. (1998). *Computer Corpus Lexicography*. Edinburgh University Press.
- Partington, A. (2004). ‘Utterly content in each other’s company’: Semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, 9(1), 131–156. <https://doi.org/10.1075/ijcl.9.1.07par>
- Pearson, J. (1996). *Terms and Text* [PhD Thesis]. University of Birmingham.
- Pearson, J. (1998). *Terms in context* (τ. 1). John Benjamins. <https://doi.org/10.1075/scl.1>
- Saridakis, I. (2013). Cross-linguistic Semantics of International Law. A corpus-informed translation of A. Cassese’s International Law into Greek. *Linguistica Antverpiensia New Series*, 12 (2013 Special Issue on “Research Models and Methods in Legal Translation”), 197–215. <https://doi.org/10.52034/lanstts.v12i.238>
- Saridakis, I., & Mouka, E. (2018). ‘Bad language’ in film translation: A corpus-driven study. *Corpora and Discourse International Conference*.

- Sifianou, M. (2003). Language variation in Greece. Στο P. Trudgill, D. Britain, & J. Cheshire (Επιμ.), *Social dialectology: In honour of Peter Trudgill* (σ. 263–273). John Benjamins.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. OUP.
- Sinclair, J. (2004). *Trust the text. Language, corpus and discourse*. Routledge.
- Sinclair, J., & Ball, J. (1996). *EAGLES preliminary recommendations on text typology. Eagles doc EAG-TCWG-ttyp/P (June 1996 version)*. <https://www.ilc.cnr.it/EAGLES96/texttyp/texttyp.html>
- Stubbs, M. (2002). *Words and Phrases. Corpus Studies of Lexical Semantics*. Blackwell.
- Teubert, W., & Čermáková, A. (2007). *Corpus Linguistics. A Short Introduction*. Continuum.
- Tognini-Bonelli, E. (2001). *Corpus Linguistics at work*. John Benjamins.
- White, P. R. R. (2015). Appraisal Theory. Στο K. Tracy, T. Sandel, & C. Ilie (Επιμ.), *The International Encyclopedia of Language and Social Interaction* (σ. 1–7). Wiley. <https://doi.org/10.1002/9781118611463.wbielsi041>
- Wodak, R. (2015). Critical Discourse Studies: A sociocognitive approach. *Discourse & Society*, 26(1), 1–13. <https://doi.org/10.1177/0957926514541355>
- Wodak, R. (2024). Critical Discourse Analysis/Studies. Στο C. Leung & J. Lewkowicz, *The Routledge Companion to English Studies* (2η έκδ., σ. 31–44). Routledge. <https://doi.org/10.4324/9781003221265-4>
- Wodak, R., & Meyer, M. (Επιμ.). (2001). *Methods of Critical Discourse Analysis*. Sage.

Νίκη Χρ. Τρύφωνα

Υποψήφια διδάκτωρ

Τμήμα Τουρκικών Σπουδών και Σύγχρονων Ασιατικών Σπουδών ΕΚΠΑ

Ηλ-ταχ.: nikitr99@icloud.com