

ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

ΣΧΟΛΗ ΑΝΘΡΩΠΙΣΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ



**ΤΜΗΜΑ ΜΕΣΟΓΕΙΑΚΩΝ ΣΠΟΥΔΩΝ & ΠΑΙΔΑΓΩΓΙΚΟ ΤΜΗΜΑ
ΔΗΜΟΤΙΚΗΣ ΕΚΠΑΙΔΕΥΣΗΣ**

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

**«ΑΝΑΛΥΣΗ ΚΑΙ ΔΙΔΑΣΚΑΛΙΑ ΠΡΩΤΗΣ ΚΑΙ ΔΕΥΤΕΡΗΣ/ΞΕΝΗΣ
ΓΛΩΣΣΑΣ»**

«Σχεδίαση και κατασκευή μαθητικού σώματος κειμένων της ελληνικής μαθητών/ μαθητριών του Γυμνασίου. Μια πρώτη επεξεργασία για τη μελέτη θεμάτων ορθογραφίας»

« Constructing a greek high school students learner corpus. A first process for the study of orthography issues»

Μεταπτυχιακή Διπλωματική Εργασία της:

Χαραλάμπους Μαρία, ΑΜ: 438218030

Επιβλέπουσα: Φραντζή Κατερίνα

Μέλη εξεταστικής επιτροπής: 1) Καραντζόλα Ελένη

2) Νικολού Καλομοίρα

3) Φραντζή Κατερίνα

ΡΟΔΟΣ, ΙΟΥΝΙΟΣ 2020

«Οι απόψεις που διατυπώνονται στην παρούσα διπλωματική εργασία εκφράζουν αποκλειστικά τη συγγραφέα»

ΠΕΡΙΛΗΨΗ

Στην παρούσα διπλωματική δημιουργείται ένα μαθητικό σώμα κειμένου επισημειωμένο όπου μέσω του πρόγραμματος AntConc3.5.8w θα γίνει επεξεργασία των ορθογραφικών λαθών από εκατόν δεκατέσσερα (114) γραπτά μαθητών/τριών Γυμνασίου, σύμφωνα με την κατηγοριοποίηση των Πρωτόπαπα και Σκαλούμπακα (2010). Πιο συγκεκριμένα, τα λάθη ταξινομούνται και αναλύονται σε τέσσερις βασικές κατηγορίες: 1) γραμματικά - μορφολογικά, 2) οπτικά – ετυμολογικά (ιστορικής ορθογραφίας) 3) φωνολογικά και 4) στην κατηγορία «άλλα λάθη» που περιλαμβάνονται α) οι λάθος ενωμένες λέξεις, β) ο λανθασμένος χωρισμός λέξεων, γ) η λάθος αντικατάσταση κεφαλαίων/μικρών γραμμάτων.

Τίθενται στη συνέχεια κάποιες υποθέσεις ως προς α) το φύλο, β) την ηλικία/τάξη και γ) την καταγωγή που θα αποτελέσουν το αντικείμενο της μελέτης μας. Πιο συγκεκριμένα:

Ως προς το φύλο: Υ1α: Τα αγόρια με Γ1 την ελληνική, κάνουν αθροιστικά περισσότερα ορθογραφικά λάθη από τα κορίτσια με Γ1 την ελληνική. Υ1β: Δεν υπάρχουν διαφοροποιήσεις φύλου ως προς την κατηγορία ορθογραφικών λαθών που γίνονται συχνότερα.

Ως προς την ηλικία/τάξη: Υ2α: Τα αγόρια με Γ1 την ελληνική, σε όλες τις τάξεις (Α', Β', Γ' Γυμνασίου) κάνουν αθροιστικά περισσότερα λάθη από τα κορίτσια. Υ2β: Όσο μεγαλώνει η τάξη (κατ'επέκταση και ηλικία), τα ορθογραφικά λάθη των αγοριών με Γ1 την ελληνική μειώνονται. Υ2γ: Αντίστοιχα, όσο μεγαλώνει η τάξη, τα ορθογραφικά λάθη των κοριτσιών με Γ1 την ελληνική μειώνονται.

Ως προς την καταγωγή: Υ3α: Οι αλλοδαποί/ές μαθητές/τριες του Γυμνασίου παράγουν την πλειονότητα των ορθογραφικών τους λαθών στις ίδιες ακριβώς κατηγορίες με τους/τις μαθητές/τριες με ελληνική καταγωγή.

Τα αποτελέσματα της έρευνας έδειξαν πως οι υποθέσεις μας μερικώς επιβεβαιώθηκαν. Πιο συγκεκριμένα οι υποθέσεις γύρω από το φύλο βγήκαν αληθείς. Τα αγόρια πράγματι έκαναν αθροιστικά περισσότερα ορθογραφικά λάθη από τα κορίτσια στο σύνολο των τριών τάξεων του Γυμνασίου, χωρίς ωστόσο η ποσοστιαία τους διαφορά να είναι ιδιαίτερα μεγάλη (10%). Επιπλέον παρατηρήθηκε πως οι δυσκολίες που αντιμετωπίζουν στο σύνολο τα αγόρια και τα

κορίτσια με Γ1 την ελληνική, εντοπίζονται στην ίδια ορθογραφική κατηγορία και πιο συγκεκριμένα ήταν λάθη ετυμολογικά - οπτικά (ή λάθη ιστορικής ορθογραφίας).

Όσον αφορά την ηλικία/τάξη, τα αγόρια με Γ1 την ελληνική και στις τρεις τάξεις του Γυμνασίου, όντως έκαναν αθροιστικά περισσότερα λάθη από τα κορίτσια. Επίσης, όσο μεγάλωναν τα αγόρια ηλικιακά και προχωρούσαν τις τάξεις, πράγματι τα ορθογραφικά τους λάθη μειώθηκαν. Ωστόσο όσο μεγάλωναν τα κορίτσια ηλικιακά και προχωρούσαν τάξη τα ορθογραφικά λάθη δεν ακολούθησαν πτωτική τάση και στις τρεις τάξεις του Γυμνασίου. Από την Α' στη Β' πράγματι υπήρξε μείωση ορθογραφικών και μάλιστα σημαντική. Στη Γ' Γυμνασίου όμως τα ορθογραφικά των κοριτσιών απέκτησαν αυξητική τάση. Επομένως, αυτή η υπόθεση δεν επιβεβαιώθηκε.

Ως προς την καταγωγή, Υ3α): Οι αλλοδαποί/ες μαθητές/τριες του Γυμνασίου πράγματι παρήγαγαν τα περισσότερα λάθη στις ίδιες κατηγορίες με τους/τις Έλληνες/ιδες μαθητές/τριες. Παρατηρήθηκαν δηλαδή λάθη κυρίως στην κατηγορία ετυμολογικά – οπτικά (ή λάθη ιστορικής ορθογραφίας), και ακολούθησαν τα γραμματικά - μορφολογικά.

ΛΕΞΕΙΣ-ΚΛΕΙΔΙΑ

Σώματα Κειμένων, Μαθητικά Σώματα Κειμένων, Γλωσσικό λάθος, Ορθογραφικό λάθος, Πρώτη γλώσσα, Δεύτερη γλώσσα

ABSTRACT

In this present thesis an annotated learner corpus is created where, via the programme AnConc3.5.8w, the spelling mistakes will be processed from 114 reports by high school students according to the categorization by Protopapa and Skaloumpaka (2010). More specifically, the spelling errors are sorted and analyzed into 4 main categories: 1) Grammatical 2) Orthographic 3) Phonological and 4) in the category “other mistakes” that include a) the wrongly joined words b) the wrong separation of words c) the wrong replacement of upper/lower case letters.

Some assumptions are set as to the a) gender, b) age/class and c) the origin that consists of the object of our study. More specifically:

As per the gender A1a: The boys with L1 in Greek, in total make more errors than girls with L1 in Greek. A1b: there is no differentiation in gender as to the spelling errors category that happens more often.

As per the age/class A2a: The boys with L1 in Greek in all classes (1st, 2nd, 3rd class of high school) make more mistakes than girls. A2b: The bigger the class (by expansion and by age) the less spelling errors are made by boys with L1 in Greek. A2c: The bigger the class the less spelling errors are made girls with L1 in Greek.

As per the origin A3a: the foreign high school students make more spelling errors in the same categories with the Greek students.

The results from the research show that our assumptions have partly been confirmed. Specifically the assumptions about the gender came out true. The boys really made more spelling errors than the girls in total of all 3 years of high school without however the difference in percentage to be especially big (10%). In addition, it was noticed that the difficulties that the boys and girls with L1 in Greek have to encounter are located in the same spelling category and specifically in the orthographic spelling errors.

Concerning the age/class, the boys with L1 in Greek in all 3 years of high school indeed made more mistakes than the girls. Also, as the boy's age and move up years, indeed their spelling mistakes decrease. However, as the girl's age and move up class the spelling mistakes do not follow a downward trend in all 3 classes of high school. In the first 2 classes of high school actually there is a decrease in spelling errors and in fact an important one. In class 3 of high school though, the girls spelling errors gain an upward trend. Therefore, that assumption was not confirmed.

As per the origin A3a, the foreign high school students indeed made more mistakes in the same categories with the Greek students. Mistakes were observed mainly in the category orthographic spelling errors then followed the grammatical spelling errors

KEY-WORDS

Text Corpora, Learner Corpora, linguistic error, orthographic error, First language, Second language.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ	3
ΛΕΞΕΙΣ-ΚΛΕΙΔΙΑ	4
ABSTRACT	5
ΒΡΑΧΥΓΡΑΦΙΕΣ	11
1. ΕΙΣΑΓΩΓΗ	13
2. ΣΩΜΑΤΑ ΚΕΙΜΕΝΩΝ	15
2.1. ΕΝΝΟΙΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ	15
2.2. ΣΩΜΑΤΩΝ ΚΕΙΜΕΝΩΝ ΚΑΙ ΑΝΤΙΠΡΟΣΩΠΕΥΤΙΚΟΤΗΤΑ	17
2.3. ΣΥΝΤΟΜΗ ΙΣΤΟΡΙΚΗ ΑΝΑΦΟΡΑ	18
2.4. ΚΑΤΗΓΟΡΙΕΣ ΣΩΜΑΤΩΝ ΚΕΙΜΕΝΩΝ	19
2.4.1. <i>Επισημειωμένα και μη Σώματα Κειμένων</i>	21
2.5. ΕΡΓΑΛΕΙΑ ΑΞΙΟΠΟΙΗΣΗΣ ΣΩΜΑΤΩΝ ΚΕΙΜΕΝΩΝ	22
2.6. ΣΩΜΑΤΑ ΚΕΙΜΕΝΩΝ ΤΗΣ ΝΕΑΣ ΕΛΛΗΝΙΚΗΣ ΓΛΩΣΣΑΣ	25
3. ΣΩΜΑΤΑ ΚΕΙΜΕΝΩΝ ΜΑΘΗΤΩΝ	28
3.1. ΟΡΙΣΜΟΣ	28
3.2. ΜΑΘΗΤΙΚΑ ΣΩΜΑΤΑ ΚΕΙΜΕΝΩΝ ΤΗΣ ΝΕΑΣ ΕΛΛΗΝΙΚΗΣ ΓΛΩΣΣΑΣ ΩΣ Γ2	28
3.3. ΜΑΘΗΤΙΚΑ ΣΩΜΑΤΑ ΚΕΙΜΕΝΩΝ ΤΗΣ ΑΓΓΛΙΚΗΣ ΓΛΩΣΣΑΣ ΩΣ Γ1	30
3.4. ΜΑΘΗΤΙΚΑ ΣΩΜΑΤΑ ΚΕΙΜΕΝΩΝ ΤΗΣ ΑΓΓΛΙΚΗΣ ΓΛΩΣΣΑΣ ΩΣ Γ2	31
4. ΒΑΣΙΚΕΣ ΕΦΑΡΜΟΓΕΣ ΤΩΝ ΣΩΜΑΤΩΝ ΚΕΙΜΕΝΩΝ ΣΤΗ ΓΛΩΣΣΙΚΗ ΔΙΔΑΣΚΑΛΙΑ	36
4.1. Η ΣΥΜΒΟΛΗ ΤΗΣ ΓΛΩΣΣΟΛΟΓΙΑΣ ΣΩΜΑΤΩΝ ΚΕΙΜΕΝΩΝ ΣΤΗ ΓΛΩΣΣΙΚΗ ΔΙΔΑΣΚΑΛΙΑ ..	36
4.2. Η ΑΡΧΗ ΤΗΣ ΓΛΩΣΣΟΔΙΔΑΚΤΙΚΗΣ ΑΞΙΟΠΟΙΗΣΗΣ ΤΩΝ ΣΩΜΑΤΩΝ ΚΕΙΜΕΝΟΥ	37
4.3. Η ΑΞΙΟΠΟΙΗΣΗ ΤΩΝ ΣΩΜΑΤΩΝ ΚΕΙΜΕΝΩΝ ΣΤΗ ΓΛΩΣΣΙΚΗ ΔΙΔΑΣΚΑΛΙΑ	38
4.4. ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΣΤΗ ΔΙΔΑΣΚΑΛΙΑ ΜΕ ΜΑΘΗΤΙΚΑ ΣΩΜΑΤΑ ΚΕΙΜΕΝΩΝ	41
4.5. ΑΞΙΟΠΟΙΗΣΗ ΜΑΘΗΤΙΚΟΥ ΣΩΜΑΤΟΣ ΚΕΙΜΕΝΩΝ ΣΤΗ ΓΛΩΣΣΙΚΗ ΔΙΔΑΣΚΑΛΙΑ	42
5. ΓΛΩΣΣΙΚΟ ΛΑΘΟΣ	45
5.1. ΟΡΙΣΜΟΣ ΚΑΙ ΠΡΟΒΛΗΜΑΤΙΣΜΟΙ.....	45
5.2. ΕΛΛΗΝΙΚΗ ΓΛΩΣΣΑ ΚΑΙ ΙΣΤΟΡΙΚΗ ΟΡΘΟΓΡΑΦΙΑ.....	46
5.3 ΚΑΤΗΓΟΡΙΕΣ ΓΛΩΣΣΙΚΩΝ ΛΑΘΩΝ	48
5.3.1. <i>Ορθογραφικά λάθη</i>	51
5.3.2 <i>Κατηγοριοποίηση Ορθογραφικών λαθών</i>	52
5.4. ΜΕΛΕΤΕΣ ΑΝΑΛΥΣΗΣ ΛΑΘΩΝ ΣΤΗΝ ΕΛΛΗΝΙΚΗ ΩΣ Γ1 ΚΑΙ Γ2	55
6. ΕΠΕΞΕΡΓΑΣΙΑ ΜΑΘΗΤΙΚΟΥ ΣΩΜΑΤΟΣ ΚΕΙΜΕΝΟΥ ΜΑΘΗΤΩΝ ΓΥΜΝΑΣΙΟΥ ΓΙΑ ΤΗΝ ΑΝΑΛΥΣΗ ΟΡΘΟΓΡΑΦΙΚΩΝ ΛΑΘΩΝ	58
6.1. ΤΑΥΤΟΤΗΤΑ ΤΗΣ ΕΡΕΥΝΑΣ.....	58
6.2. ΣΚΟΠΟΣ ΚΑΙ ΣΤΟΧΟΙ	59
6.3. ΥΠΟΘΕΣΕΙΣ ΤΗΣ ΕΡΕΥΝΑΣ	61

6.4. ΜΟΡΦΕΣ ΑΝΑΛΥΣΗΣ.....	61
6.5. ΜΕΘΟΔΟΛΟΓΙΑ	62
6.6. ΜΑΘΗΤΙΚΟ ΣΩΜΑ ΚΕΙΜΕΝΩΝ	62
6.7. ΚΡΙΤΗΡΙΑ ΕΠΙΛΟΓΗΣ ΤΟΥ ΔΕΙΓΜΑΤΟΣ.....	63
6.8. ΔΙΑΔΙΚΑΣΙΑ ΔΕΙΓΜΑΤΟΛΗΨΙΑΣ.....	63
6.9. ΠΕΡΙΓΡΑΦΗ ΤΕΛΙΚΟΥ ΔΕΙΓΜΑΤΟΣ.....	68
6.10. ΠΕΡΙΟΡΙΣΜΟΙ ΤΗΣ ΕΡΕΥΝΑΣ.....	72
ΑΠΟΤΕΛΕΣΜΑΤΑ	74
ΣΥΜΠΕΡΑΣΜΑΤΑ.....	80
ΠΡΟΕΚΤΑΣΕΙΣ	85
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	87
ΠΑΡΑΡΤΗΜΑ.....	103

ΕΥΧΑΡΙΣΤΙΕΣ

Ευχαριστώ θερμά την καθηγήτρια Πληροφορικής και Πρόεδρο του Τμήματος Μεσογειακών Σπουδών του Πανεπιστημίου Αιγαίου, κα Κατερίνα Φραντζή, που δέχτηκε να αναλάβει την εποπτεία μου για την ολοκλήρωση της διπλωματικής μου εργασίας, την καθοδήγηση που μου προσέφερε για την επιλογή του θέματος και τη διαχείριση του υλικού, καθώς και την ελευθερία που μου παραχώρησε για την επεξεργασία των δεδομένων.

Επίσης, θα ήθελα να ευχαριστήσω τις καθηγήτριες γλωσσολογίας κα. Καραντζόλα Ελένη και την κα Νικολού Καλομοίρα του Τμήματος Μεσογειακών Σπουδών, για το χρόνο που αφιέρωσαν να μελετήσουν τη διπλωματική μου εργασία και να αποτελέσουν μέλη της εξεταστικής μου επιτροπής.

Στους γονείς μου...

ΒΡΑΧΥΓΡΑΦΙΕΣ

Ελληνικές:

ΣΚ: Σώματα Κειμένων

ΗΣΚ: Ηλεκτρονικό Σώμα Κειμένων

ΜΣΚ: Μαθητικά Σώματα Κειμένων

ΕΘΕΓ : Εθνικός Θησαυρός της Ελληνικής γλώσσας

ΣΕΚ: Σώμα Ελληνικών Κειμένων

ΚΕΓ: Κέντρο Ελληνικής Γλώσσας

Α.Π.Θ: Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

ΕΚΠΑ: Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Γ1: Πρώτη γλώσσα

Γ2: Δεύτερη γλώσσα

ΤΠΕ: Τεχνολογίες Πληροφορίας και επικοινωνιών

ΕΣΚΕΙΜΑΘ: Εκπαίδευση Αλλοδαπών και Παλινοστούντων Μαθητών

ΣΕΠΑΜΕ2: Σώμα Εξελικτικών Παραγωγών Μαθητών της Ελληνικής ως Γ2

Ξένοι:

LOB:	Lancaster- Oslo/ Bergen
BNC:	British National Corpus
ICE:	International Corpus of English
CHILDES:	Child Language Data Exchange System
LCPW:	Lancaster Corpus of Children's Project Writing
POW:	Polytechnic of Wales Corpus
ICLE:	International Corpus of English
LINDSEI:	Louvain International Database of Spoken English Interlanguage
LOCNEC:	Louvain Corpus Native English Conversation
ICNALE:	International Corpus Network of Asian Learners of English
CLEC:	Chinese Learner English Corpus
SWECCL:	The Spoken & Written English Corpus of Chinese Learners
SLC:	Spoken Learner Corpus
LANCAWE:	Lancaster Corpus of Academic Written English
MELD:	Montclair Electronic Language Learners' Database
VOICE:	Vienna-Oxford International Corpus of English
HKUST:	Hong Kong University of Science & Technology Corpus
TLEC:	Thai English Learner Corpus
CLC:	Cambridge Learner Corpus

1. ΕΙΣΑΓΩΓΗ

Για τη γλωσσολογική έρευνα, η αξία ενός Σώματος Κειμένων έγκειται στη δυνατότητα που προσφέρει στον/στην ερευνητή/τρια να του/της παρέχει πρόσβαση σε μεγάλο όγκο γλωσσικού υλικού, το οποίο συγκεντρώθηκε με συγκεκριμένα κριτήρια. Σκοπός της δημιουργίας του είναι να αποτελέσει αντιπροσωπευτικό υλικό της γλώσσας αλλά και το εργαλείο της επεξεργασίας αυτών των δεδομένων. Μια κατηγορία πιο ειδική αποτελούν τα Σώματα Κειμένων Μαθητικού λόγου, όπου βασικό τους πλεονέκτημα είναι η συγκέντρωση αυθεντικού λόγου των μαθητών/τριών μιας γλώσσας. Αυτό αποτελεί πολύτιμη βοήθεια για τον/την εκπαιδευτικό ή τον/την ερευνητή/τρια καθώς προσφέρει τη δυνατότητα εξαγωγής συμπερασμάτων και άντλησης παραδειγμάτων μέσα από τη μελέτη της ίδιας της γλώσσας των μαθητών/τριών, χωρίς να στηρίζονται σε νοητά παραδείγματα και εικασίες.

Επιπρόσθετα, αν σκεφτεί κανείς πως ένα από τα σημαντικότερα παιδαγωγικά προβλήματα στο σχολείο, τόσο στην Ελλάδα όσο και στο εξωτερικό αποτελεί το ορθογραφικό ζήτημα και αν αναλογιστεί ταυτόχρονα, πόσο μεγάλη και συστηματική προσπάθεια επιδέχεται η γλωσσική παιδεία όλα αυτά τα χρόνια, θα απογοητευτεί από τα αποτελέσματα που σημειώνονται από μελετητές/τριες της γλώσσας, εκπαιδευτικούς και γονείς. Τα ερωτήματα που δημιουργούνται και απασχολούν τους ειδικούς είναι πολλά. Κυρίως αυτών που σχεδιάζουν τα αναλυτικά προγράμματα και ψάχνουν τρόπους για την κατάκτηση της ορθογραφικής δεξιότητας.

Τις τελευταίες δεκαετίες λόγω της αυξημένης προσέλευσης αλλοδαπών στη χώρα μας, το θέμα της ένταξης των παιδιών με μεταναστευτική ταυτότητα στα ελληνικά σχολεία κάνει το θέμα του γλωσσικού λάθους πιο επίκαιρο από ποτέ. Μια σειρά από μελέτες προσπαθούν να εντοπίσουν και να ερμηνεύσουν τα λάθη αυτά ώστε να βοηθήσουν στην έγκαιρη αντιμετώπιση τους. Μια σύγχρονη προσέγγιση στο θέμα αυτό επιτυγχάνεται μέσα από τη δημιουργία επισημειωμένων ή μη Μαθητικών Σωμάτων Κειμένων που σε πρώτο επίπεδο επιχειρεί να εντοπίσει και να κατηγοριοποιήσει τα λάθη των μαθητών/τριών ώστε να μπορέσει να τα ερμηνεύσει αλλά και σε ένα δεύτερο επίπεδο, να μπορέσουν τα γραπτά αυτά να αξιοποιηθούν για εκπαιδευτικό υλικό σε μια προσπάθεια βελτίωσης των αδυναμιών τους.

Οι σύγχρονες μελέτες ωστόσο που αφορούν τα γλωσσικά λάθη των μαθητών και μαθητριών στην ελληνική ως πρώτη γλώσσα (Γ1) δεν είναι ιδιαίτερα εκτενείς και κυρίως αυτές που

αφορούν τους/τις εφήβους. Οι μελέτες ως επί το πλείστον εστιάζουν στην παραγωγή γραπτού λόγου από μαθητές και μαθήτριες με δεύτερη γλώσσα (Γ2) την ελληνική. Προσπαθούν να εντοπίσουν τις αδυναμίες των δίγλωσσων ώστε να μπορέσουν να έχουν μια αποτελεσματικότερη παιδαγωγική προσέγγιση απέναντι τους. Στη λογική αυτή στηριζόμενοι/ες, Έλληνες/ίδες μελετητές/τριες κατασκεύασαν μαθητικά σώματα από γραπτές παραγωγές δίγλωσσων μαθητών/τριών ώστε να αποτελέσουν εργαλείο μελέτης και επεξεργασίας τόσο για επιστημονικούς όσο και για εκπαιδευτικούς σκοπούς. Αποτελέσματα ερευνών επιβεβαίωσαν πως τα Σώματα Κειμένων καθώς και τα πιο εξειδικευμένα σώματα όπως τα μαθητικά, μπορούν να αξιοποιηθούν στη γλωσσική διδασκαλία, ενώ ταυτόχρονα μπορούν να χρησιμοποιηθούν και στην ανάλυση μαθητικού λόγου με θετικά αποτελέσματα.

Αφορμή όσων προαναφέρθηκαν ήταν η δημιουργία της παρούσας διπλωματικής εργασίας που σκοπό έχει να μελετήσει τα Σώματα Κειμένων και ειδικότερα τα Μαθητικά Σώματα, να δημιουργήσει ένα τέτοιο επισημειωμένο μαθητικό σώμα και να μελετήσει τα ορθογραφικά λάθη μαθητών και μαθητριών Γυμνασίου.

Πιο συγκεκριμένα, στην πρώτη ενότητα γίνεται αναφορά στα Σώματα Κειμένων που δημιουργήθηκαν στην ελληνική γλώσσα, εξειδικευμένα και μη.

Στη δεύτερη ενότητα γίνεται εστίαση σε όλα τα Μαθητικά Σώματα Κειμένων (στο εξής ΜΣΚ) που δημιουργήθηκαν μέχρι σήμερα για την ελληνική γλώσσα από δίγλωσσους/ες μαθητές/τριες αλλά και στα Μαθητικά Σώματα που δημιουργήθηκαν με Γ1 και Γ2 την αγγλική γλώσσα διεθνώς.

Στην Τρίτη ενότητα γίνεται αναφορά στα πλεονεκτήματα που προσφέρει η διδασκαλία με χρήση Σωμάτων Κειμένων (ΣΚ) καθώς και στην αξιοποίηση που μπορούν να έχουν στη γλωσσική διδασκαλία τα ΣΚ γενικά αλλά και ειδικότερα τα Μαθητικά Σώματα Κειμένων.

Στην τέταρτη ενότητα θα γίνει αποσαφήνιση του γλωσσικού λάθους και αναφορά στις κατηγοριοποιήσεις που έχουν προταθεί ανά καιρούς από ερευνητές/τριες. Θα εστιάσουμε ωστόσο την προσοχή μας στο ορθογραφικό λάθος που άπτεται του θέματος μας και θα αναφερθούμε σε προγενέστερες μελέτες και τα αποτελέσματά τους.

Έπειτα, στην πέμπτη και τελευταία ενότητα, μέσα από την πιλοτική κατασκευή ενός Μαθητικού Σώματος Κειμένων με πλειοψηφία γραπτών από μαθητές και μαθήτριες με Γ1 την ελληνική και λιγότερους στο ποσοστό από μαθητές και μαθήτριες με Γ2 την ελληνική,

θα αναλυθούν τα ορθογραφικά λάθη των μαθητών/τριών που παρήχθησαν από μη τυπικές μορφές εκπαίδευσης (φροντιστήρια / οικοδιδασκαλεία) από το νησί της Ρόδου.

Ειδικότερα επιχειρείται σε ένα πρώτο στάδιο η συλλογή, η οργάνωση και η ταξινόμηση των γραπτών που συγκεντρώθηκαν ώστε να αναπτυχθεί ένα όσο είναι δυνατόν αντιπροσωπευτικό επισημειωμένο Μαθητικό Σώμα Κειμένων (ΜΣΚ). Το λογισμικό που βοήθησε σε αυτό ήταν το AntConc 3.5.8w (2019), ένα ημιαυτόματο εργαλείο επεξεργασίας οποιουδήποτε Σώματος Κειμένου. Με αυτόν τον τρόπο δημιουργήθηκε ένα ΜΣΚ της Νέας Ελληνικής από γραπτά μαθητών και μαθητριών Γυμνασίου που αποτελείται από 23.300 περίπου λέξεις.

2. ΣΩΜΑΤΑ ΚΕΙΜΕΝΩΝ

2.1. Εννοιολογική προσέγγιση

Πριν δώσουμε ορισμό για τα Σώματα Κειμένων θα ήταν ορθότερο να προσδιορίσουμε τον όρο τους στενότερα και να τα εντάξουμε σε μια ευρεία σημασιολογική κατηγορία, αυτή της συλλογής κειμένων (text collection). Οι Atkins, Clear και Oster (1992, 1) διακρίνουν τέσσερα (4) είδη από συλλογές κειμένων:

- i) *Αρχείο (archive)*: Αποθετήριο κειμένων σε ηλεκτρονική μορφή χωρίς κάποια οργανωμένη διασύνδεση μεταξύ τους.
- ii) *Βιβλιοθήκη Ηλεκτρονικών Κειμένων (electronic text library)*: Συλλογή κειμένων ηλεκτρονικής μορφής με ενιαία τυποποίηση και ορισμένες συμβάσεις του περιεχομένου τους, χωρίς ωστόσο κάποια σοβαρά περιορισμένη επιλογή.
- iii) *Ηλεκτρονικό Σώμα Κειμένων (corpus)*: Πρόκειται για υποσύνολο της βιβλιοθήκης ηλεκτρονικών κειμένων με τη μόνη διαφορά πως η σχεδίαση της γίνεται με συγκεκριμένες προδιαγραφές και για συγκεκριμένους σκοπούς.
- iv) *Υποσώμα Κειμένων (subcorpus)*: Πρόκειται για υποσύνολο του Σώματος Κειμένων που αποτελείται είτε από στατικά στοιχεία ενός σύνθετου σώματος κειμένων είτε από μια δυναμικά προσδιορισμένη επιλογή ενός σώματος κειμένων για να πραγματοποιηθεί μια συγκεκριμένη ανάλυση.

Έχοντας κάνει τις παραπάνω διευκρινίσεις μένει να προσθέσουμε πως τα Σώματα Κειμένων (στο εξής ΣΚ) αποτελούν θεμελιώδες εργαλείο έρευνας για τον κλάδο της γλωσσολογίας. Κατά καιρούς έχουν δοθεί διάφοροι ορισμοί γι' αυτά.

Ο Sinclair (1996, 27), όρισε τα ΣΚ ως συλλογές από γλωσσικά στοιχεία που επιλέχθηκαν με ορισμένα γλωσσολογικά κριτήρια για να αποτελέσουν ένα αντιπροσωπευτικό γλωσσικό δείγμα.

Ο Kennedy (1998, 14) όρισε τα ΣΚ ως δομημένες συλλογές κειμένων, που δημιουργήθηκαν για να αναλυθούν γλωσσολογικά και να απαρτίσουν αντιπροσωπευτικό δείγμα της γλώσσας. Τονίζει πως τα ΣΚ μπορεί να αποτελούνται από ολόκληρα κείμενα ή συλλογές ολόκληρων κειμένων προφορικού ή γραπτού λόγου. Μπορεί ακόμη να αποτελούνται από δείγματα συνεχών κειμένων που λαμβάνονται από ολόκληρα κειμενικά συγγράμματα, είτε ακόμη, από συλλογές παραπομπών. Το βέβαιο είναι πως η δημιουργία τους δεν πληροί αποκλειστικά και μόνον ένα σκοπό. Ο σκοπός αυτός είναι που θα επηρεάσει το σχεδιασμό, το μέγεθος και τη φύση του κάθε ΣΚ.

Ο Sampson (2001, 6) παρόμοια, ορίζει τα ΣΚ «ως εκτεταμένο δείγμα αυθεντικής χρήσης της υπό εξέταση γλώσσας, που συγκροτείται και χρησιμοποιείται ως πηγή στοιχείων για την παραγωγή ή εξέταση υποθέσεων για τη φύση της γλώσσας»

Ο McEnery (2005, 449), αναφέρει πως ως ΣΚ θα πρέπει να ορίζεται μόνο μια σωστά δομημένη συλλογή δεδομένων που συγκεντρώθηκε σε ένα πλαίσιο δειγματοληψίας και είναι ειδικά σχεδιασμένη για να επιτρέπει τη μελέτη σε συγκεκριμένα γλωσσικά στοιχεία, αναλόγως των δεδομένων που συλλέχθηκαν.

Αντίστοιχα και η Bennett (2010, 12), τα όρισε ως μια συλλογή αρχών με αυθεντικά κείμενα αποθηκευμένα σε ηλεκτρονική μορφή, τα οποία μπορούν να χρησιμοποιηθούν για να ανακαλύψουν πληροφορίες σχετικά με τη γλώσσα, οι οποίες μπορεί να μην έχουν παρατηρηθεί μόνο μέσω της διαίσθησης. Τα Σώματα Κειμένων διαιρούνται σε περίπου οχτώ κατηγορίες, τα γενικευμένα (generalized), τα εξειδικευμένα (specialized), τα μαθητικά (learner), τα παιδαγωγικά (pedagogic), τα ιστορικά (historical), τα συγκρίσιμα (comparable), τα σώματα παραλληλισμού (parallel) και τα υπό παρακολούθησης (monitor).

Συμπυκνώνοντας, θα υποστηρίζαμε πως τα ΣΚ είναι ένα σύνολο από γλωσσικά δεδομένα που συγκεντρώνονται με βάση κάποια συγκεκριμένα κριτήρια και αποθηκεύονται σε ηλεκτρονική μορφή ώστε εύκολα και γρήγορα να μπορούν να ανακαλούνται.

2.2. Σωμάτων Κειμένων και αντιπροσωπευτικότητα

Ο Sinclair (1996), μέσα από τον ορισμό που δίνει για τα ΣΚ είναι ο εισηγητής του όρου *αντιπροσωπευτικότητα* (representativeness). Κρίνεται αναγκαίο για οποιοδήποτε ΣΚ να προσφέρει μια εικόνα που να είναι αντιπροσωπευτική της γλώσσας είτε της γλωσσικής ποικιλίας που συλλέγεται. Παρόλο που δεν είναι πάντα θεμιτό να τίθενται όρια και παράμετροι με τα οποία θα συλλέγεται συγκεκριμένος αριθμός και συγκεκριμένο κειμενικό είδος, αυτό που έχει μέγιστη σημασία όταν αναφερόμαστε στον όρο αντιπροσωπευτικότητα είναι η έμφαση στην ποικιλομορφία των κειμένων καθώς και των κειμενικών ειδών, στην ύπαρξη και αντιστάθμιση των μεταξύ τους αναλογιών και ταυτόχρονα στην αυθεντικότητα όλων των δεδομένων που συγκεντρώνονται. Επιπλέον, από ένα ΣΚ που συγκροτεί αντιπροσωπευτικό δείγμα, μπορεί να προκύψουν είτε γενικότερα συμπεράσματα για τη γλώσσα είτε για τη γλωσσική ποικιλία που περιέχεται. Ακόμη, ο Sampson στον ορισμό του, επισημαίνει πως ένα αποθηκευμένο σώμα κειμένων δεν είναι δυνατόν να παράγει από μόνο του πληροφορίες για τη γλώσσα, αλλά αποτελεί τη βάση για τη δημιουργία υποθέσεων από τον/την ερευνητή/τρια, ώστε να οδηγήσει σε εξαγωγή επαληθευμένων συμπερασμάτων. Επομένως, η παρέμβαση του/της γλωσσολόγου είναι απαραίτητη ώστε να παραχθούν ερωτήματα αλλά και να αξιολογηθούν τα αποτελέσματα που προκύπτουν (Γούτσος και Φραγκάκη 2015, 41).

Πιο αναλυτικά μέσα από τα ΣΚ μπορούμε να εξάγουμε γλωσσικά δεδομένα που έχουν τα παρακάτω χαρακτηριστικά:

Εμπειρικά: δεν στηρίζονται σε εικασίες ή την αντίληψη του/της μελετητή/τριας αλλά προκύπτουν από την ίδια την πραγματικότητα.

Αυθεντικά: δεν αντλούν στοιχεία από πειράματα ή άλλων ειδών τεχνικές συνθήκες αλλά από τη φυσική προφορική ή γραπτή παραγωγή λόγου των ομιλητών/τριών της γλώσσας.

Συστηματικά: έχουν συγκεντρωθεί βάσει συγκεκριμένων κριτηρίων και αρχών και όχι με αυθαίρετο τρόπο.

Κειμενικά: περιέχουν πλήρη κείμενα ή τμήματα κειμένων χωρίς τον περιορισμό μεμονωμένων λέξεων και προτάσεων.

Εκτεταμένα: έχουν μεγάλο όγκο και δυνατότητα παρουσίασης πολλών παραδειγμάτων. Όπως αναφέρει και ο Sinclair (1991,100), όταν μελετάς ένα μεγάλο κομμάτι της γλώσσας παράλληλα η γλώσσα δείχνει πολύ διαφορετική (Γούτσος και Φραγκάκη 2015, 15).

2.3. Σύντομη Ιστορική Αναφορά

Στις αρχές της δεκαετίας του 1960 δημιουργήθηκε το πρώτο Ηλεκτρονικό Σώμα Κειμένων (ΗΣΚ) το Brown Corpus, με πάνω από ένα εκατομμύριο λέξεις σε κείμενα διαφόρων ειδών λόγου (πεντακόσια δείγματα με πάνω από δύο χιλιάδες λέξεις το καθένα) που είναι δημοσιευμένα το 1963-1964 και γραμμένα στα Αγγλικά των Η.Π.Α. Μια δεκαετία αργότερα όμως μεταξύ 1970-1976, δημιουργήθηκε η Βρετανική εκδοχή του και ονομάστηκε Lancaster-Oslo/Bergen (LOB Corpus). Αυτά τα δύο μαζί βοήθησαν στο να πραγματοποιηθούν γλωσσικές και πολιτισμικές συγκρίσεις μεταξύ Βρετανικών και Αμερικάνικων Αγγλικών (Sketch Engine χ.χ.).

Παρόλο που στο πέρασμα των χρόνων έχουν δημιουργηθεί πολλά ΣΚ, από τα πιο γνωστά στην αγγλική γλώσσα είναι το British National Corpus (BNC) με άνω από εκατό εκατομμύρια λέξεις, το Bank of English με περιεχόμενο 450 εκατομμυρίων λέξεων και το International Corpus of English (ICE) με περιεχόμενο είκοσι εκατομμυρίων λέξεων (Braun 2005, 48). Αυτά τα ΣΚ ήταν η βάση για πλήθος γλωσσικών ερευνών και αναφορών όπως βιβλίων γραμματικής, λεξικά κ.α. Στις μέρες μας τέτοιου είδους ΣΚ ξεπερνούν τις ένα δισεκατομμύριο λέξεις (πχ. Cambridge International Corpus, Oxford English Corpus). Αντίστοιχα, παρουσιάστηκε μεγάλη ανάπτυξη δημιουργίας ΣΚ και σε άλλες γλώσσες αλλά με μεγαλύτερη χρονική καθυστέρηση (Γούτσος 2006).

Στην Ελλάδα, τα πρώτα ΣΚ σε ηλεκτρονική μορφή έκαναν την εμφάνισή τους τη δεκαετία του 1980 και περιελάμβαναν παλαιότερα έργα της λογοτεχνίας (πχ. απομνημονεύματα του Μακρυγιάννη, Κρητική λογοτεχνία), ενώ ο Goutsos κ.α (1994 στο Γούτσος 2015, 48) διαπιστώνουν πως παρά το ότι υπάρχουν σχετικά ερευνητικά προγράμματα, τα Σώματα Κειμένων, είτε χρησιμοποιούνται μερικώς στη γλωσσική έρευνα, είτε δε χρησιμοποιούνται καθόλου.

2.4. Κατηγορίες Σωμάτων Κειμένων

Ανάλογα τις προδιαγραφές σχεδίασης και τη χρήση για την οποία προορίζονται τα ΣΚ διαφέρουν. Μια ενδεικτική κατηγοριοποίηση που αναφέρουν οι Μικρός και Φουντοπούλου (χ.χ., 2-4) αφορά:

Το επίπεδο λόγου: Τα ΣΚ ανάλογα την πηγή των δεδομένων τους χαρακτηρίζονται σε προφορικά και γραπτά. Τα πρώτα ΣΚ που δημιουργήθηκαν ήταν κυρίως γραπτά (πχ. Brown Corpus), ωστόσο σύντομα δημιουργήθηκαν και προφορικά ΣΚ (πχ. London-Lund Corpus, Wellington Corpus of Spoken New Zealand English). Τα πιο επίκαιρα είναι τα μεικτά ΣΚ (πχ. British National Corpus), που με χρήση προφορικών δεδομένων καλύπτουν συγκεκριμένο ποσοστό από αυτά έναντι των γραπτών.

Την εθνικότητα: Τα ΣΚ μπορούν να κατηγοριοποιηθούν και ανάλογα την εθνική γλωσσική ποικιλία που καλύπτουν. Αυτό επιτρέπει μια συγκριτική μελέτη του τρόπου που η ίδια γλώσσα μεταβάλλεται στο πλαίσιο μιας διαφορετικής εθνότητας. (πχ. Corpus de referencia del Español actual που εμπεριέχει υποσώματα από κράτη εκτός Ισπανίας που κάνουν χρήση της ισπανικής γλώσσας).

Την ιστορική διάσταση: Ανάλογα με τη χρονολογία που έχουν τα κείμενα που περιλαμβάνονται σε ένα ΣΚ θέτονται και τα αντίστοιχα ερευνητικά ερωτήματα. Γι' αυτό το λόγο υπάρχουν ΣΚ που αποτελούνται από κείμενα που καλύπτουν κάποια συγκεκριμένη χρονική στιγμή είτε σε ιστορικό είτε σε συγχρονικό επίπεδο, είτε ακόμη και σε κάποια συγκεκριμένη χρονική περίοδο.

Τη διαλεκτική/κοινωνιογλωσσολογική ποικιλία: Η γλώσσα διαφοροποιείται ανάλογα τη γεωγραφική και κοινωνική της εμβέλεια. Μελέτες στην κοινωνιογλωσσολογία και τη διαλεκτολογία έκαναν χρήση αυθεντικών κειμένων από πολύ νωρίς και έτσι διαμορφώθηκε μια σχολή εμπειρικής προσέγγισης στην έρευνα γλωσσικών ποικιλιών με βασικό εκπρόσωπο τον Labov. Μια από τις πιο σημαντικές εφαρμογές στην έρευνα των ΣΚ είναι η ανάλυση της γλωσσικής ποικιλίας. Η κοινωνιογλωσσολογία και η διαλεκτολογία κάνουν χρήση των ΣΚ αυτών με αποτέλεσμα η έρευνα στους τομείς αυτούς να κινείται με γρήγορους ρυθμούς (Strassel και Cieri 2001). Τα ΣΚ επειδή πρέπει να αποτελούν αντιπροσωπευτικό δείγμα μιας γλώσσας, αποτελούνται από διαλεκτικού λόγου δείγματα και από ομιλητές και ομιλήτριες που μπορούν να παράγουν διακριτές γλωσσικές ποικιλίες. Για το λόγο αυτό έχουν δημιουργηθεί εξειδικευμένα ΣΚ ώστε να είναι εφικτή η πληρέστερη και αναλυτικότερη

μελέτη τους που σε συνδυασμό με τα συστήματα γεωγραφικής πληροφορίας (geographical information systems-GIS), σχηματίζουν ψηφιακές διαλεκτικές εγκυκλοπαίδειες (Kretzschmar και Lee, 1993).

Την ηλικία: Η ηλικία των ομιλητών/τριών που περιλαμβάνονται στο ΣΚ μπορεί να παίζει καθοριστικό ρόλο για την εξειδίκευσή του ως προς το σκοπό και το είδος της γλώσσας που συλλέχτηκε. Ένας διαχωρισμός που γίνεται, διακρίνει τα ΣΚ σε ενηλίκων και παιδιών. Ο παιδικός λόγος εάν εξεταστεί με ποσοτικά δεδομένα σε ΣΚ μπορεί να φέρει στην επιφάνεια σημαντικές πληροφορίες για τον τρόπο ανάπτυξης του γλωσσικού συστήματος σε ένα παιδί καθώς και την ποικιλία που εμφανίζει σε όλο το γραμματικό φάσμα (learner corpora) (Andersen 1990, 57).

Την ανανεωσιμότητα: Υπάρχουν ΣΚ που δύναται όταν ολοκληρώσουν τον σκοπό δημιουργίας τους να παγιωθούν ως μια κλειστή συλλογή (όπως συνέβη και στην περίπτωση του British National Corpus). Ωστόσο, μπορούν να παραμείνουν και ανοιχτά με σκοπό τη συνεχή ανανέωση και τον εμπλουτισμό τους. Η ανοιχτή σχεδίαση αποτελεί πλεονέκτημα όταν πρόκειται να χρησιμοποιηθεί για λεξικογραφική μελέτη, γιατί η χρήση λεξιλογίου αλλάζει σημαντικά σε σύντομα χρονικά διαστήματα. Αυτά τα ΣΚ ονομάζονται σώματα ελέγχου (monitor corpora), και οργανώνουν τα γλωσσικά δεδομένα σε διάφορα γλωσσικά διαστήματα επιτρέποντας την άμεση συσχέτιση της γλωσσικής παραγωγής. Ενώ λοιπόν, στα κλειστά ΣΚ κάνουν χρήση ως βασικό ποσοτικό χαρακτηριστικό τους τον αριθμό των λέξεων, σε ένα ανοιχτό ΣΚ μελετώνται πόσες λέξεις εισάγονται σε ορισμένο χρόνο (μέρα/μήνας/χρόνος).

Τη διαθεσιμότητα: Η διαθεσιμότητα σε ένα ΣΚ αφορά τόσο την οικονομική εκμετάλλευση (εμπορική, δωρεάν διάθεση), όσο και τον φορέα που διατίθεται (διαδίκτυο, διακομιστές ftp κ.α.).

Το χαρακτηρισμό: Υπάρχουν ΣΚ που δεν έχουν ειδικό χαρακτηρισμό. Όμως όταν ένα ΣΚ έχει χαρακτηρισμό αυτόματα γίνεται και περισσότερο χρηστικό πχ. γραμματικός χαρακτηρισμός των μερών του λόγου, αναλυτικός μορφολογικός χαρακτηρισμός, συντακτικός, σημασιολογικός κ.α.

Το εύρος γλωσσικής κάλυψης: Ανάλογα το εύρος και την αντιπροσωπευτικότητα που έχει ένα σώμα σε κειμενικά είδη χαρακτηρίζεται είτε ΣΚ αναφοράς (reference corpora) είτε ΣΚ ειδικής χρήσης (special corpora). Το πρώτο σχεδιάστηκε με σκοπό να δίνει αναλυτικές πληροφορίες για κάποια συγκεκριμένη γλώσσα, ενώ το δεύτερο αποτελείται από κείμενα που

έχουν κάποια γλωσσικά χαρακτηριστικά τους *επισημειωμένα* (annotated/tagged). Σε λεξιλογικό επίπεδο τα ΣΚ ειδικής χρήσης περιέχουν κείμενα με συγκεκριμένο θέμα και υψηλή συχνότητα εμφάνισης όρων που συσχετίζονται με αυτό.

Τον αριθμό των γλωσσών: Ανάλογα με τον αριθμό των γλωσσών που περιλαμβάνονται σε ένα ΣΚ μπορεί να υπάρξει και η ανάλογη κατηγοριοποίηση. Υπάρχουν τα μονόγλωσσα που περιλαμβάνουν μία μόνο γλώσσα ή διάλεκτο και τα πολύγλωσσα που έχουν κείμενα γραμμένα σε περισσότερες από μία γλώσσες. Μια πιο ειδική κατηγορία αποτελούν τα παράλληλα ΣΚ. Πρόκειται για συλλογές κειμένων μαζί με τις μεταφράσεις τους σε μία ή παραπάνω γλώσσες. Η πιο απλή μορφή είναι όταν στο ΣΚ υπάρχουν δύο μόνο γλώσσες όπου η μία αποτελεί μετάφραση των κειμένων της άλλης. Κάποιες φορές η κατεύθυνση ποικίλει. Μπορεί δηλαδή το κείμενο στην Α γλώσσα να μεταφραστεί στη γλώσσα Β, και κάποια κείμενα Β να μεταφραστούν στην Α. Τα παράλληλα ΣΚ παίζουν ενεργό ρόλο στη γλωσσική διδασκαλία γιατί βοηθούν σημαντικά στην εκ βάθους μελέτη της μετάφρασης, στην ανάπτυξη πολύγλωσσων λεξικών, στη μελέτη απόδοσης της ορολογίας και τέλος στη δημιουργία κατάλληλων εργαλείων στατιστικής μηχανικής μετάφρασης.

2.4.1. Επισημειωμένα και μη Σώματα Κειμένων

Ο όρος *επισημείωση* είναι γενικευμένος και αναφέρεται στην πρακτική πρόσθεση και ανάθεση κάποιας γλωσσολογικής πληροφορίας σε ένα σώμα κειμένου. Συχνά εξαιτίας της πολυπλοκότητας των συντακτικών δομών και της ποικιλίας των γραμματικών κατηγοριών που μπορεί να ανήκει μια λέξη, είτε πρόκειται για λέξεις και εκφράσεις με σημασιολογικό πλούτο είτε οποιονδήποτε άλλων ασαφειών στη γλώσσα μπορεί να προκύψουν δυσκολίες (σε κάποιες γλώσσες περισσότερο από κάποιες άλλες κατά την επεξεργασία των ΣΚ). Με την επισημείωση μετατρέπεται η έμμεση και άρρητη πληροφορία σε άμεση και ρητή με τη χρήση ετικετών σήμανσης. Τα επισημειωμένα σώματα ουσιαστικά εμπεριέχουν κάποια γλωσσολογική επισημείωση. Ως παράδειγμα πιο συχνών επισημειώσεων αναφέρουμε την επισημείωση στα μέρη του λόγου (ουσιαστικό, επίθετο, ρήμα κ.α) και την επισημείωση σε συντακτικές πληροφορίες (υποκείμενο, ρήμα, αντικείμενο κ.α) (Φραντζή 2012, 22-23 · Τάντος 2015,214).

Ένα ΣΚ στο οποίο δεν έχει γίνει καμία γλωσσολογική άλλη επεξεργασία λέγεται *μη επισημειωμένο* (un-annotated). Αν και σε άλλες δεκαετίες τα μη επισημειωμένα σώματα

αποδείχθηκαν χρήσιμα σε μελέτες συχνοτήτων και συσχέτισης λέξεων ή φράσεων σε διαφορετικού είδους περιβάλλοντα (συμφραζόμενα), με τη χρήση της επισημείωσης επιτρέπονται στοχευμένες αναζητήσεις και αναλύσεις που σε άλλη περίπτωση δε θα μπορούσαν να εξάγουν τη σωστή πληροφορία. (Φραντζή 2012, 22-23 · Τάντος 2015, 214).

Η επισημείωση γενικότερα σε ένα ΣΚ μπορεί να αφορά:

1. Επισημείωση των μερών του λόγου (ουσιαστικό, επίθετο κα.)
2. Επισημείωση συντακτικών δομών (υποκείμενο, ρήμα κα.)
3. Φωνητική επισημείωση (τονικότητα, παύσεις κα.)
4. Πραγματολογική επισημείωση (λεκτικές πράξεις πχ. αποδοχή, επιθυμία)
5. Επισημείωση σημασιολογικών όρων (κατηγοριοποίηση λέξεων ή φράσεων)
6. Στυλιστική επισημείωση (πχ. ευθύς/πλάγιος λόγος)
7. Επισημείωση κειμένου (πχ. υπερπροτασικές αναφορικές συνδέσεις αντωνυμιών)
8. Επισημείωση στα λάθη των μαθητών/τριών (πχ.μαρκάρισμα ορθογραφικών λαθών)

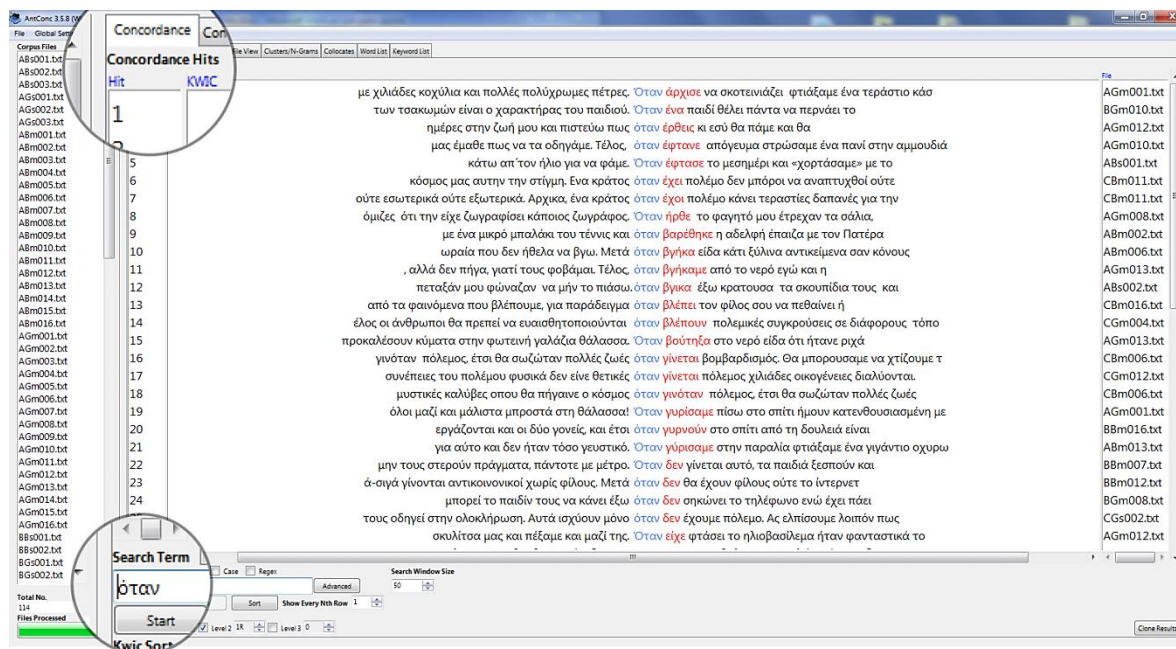
Για να μπορέσει να επιτευχθεί η επισημείωση στις άνω κατηγορίες ωστόσο θα πρέπει ο μελετητής/τρια να γνωρίζει τη γραμματική της υπό εξέταση φυσικής γλώσσας γιατί η επισημείωση με αυτόματο τρόπο από τον υπολογιστή δεν είναι εφικτή. Κρίνεται απαραίτητο μια ομάδα από γλωσσολόγους να αποτελέσουν το ενδιάμεσο στάδιο της επισημείωσης πριν ξεκινήσει η έρευνα. Η δημιουργία επομένως γλωσσικών πόρων είναι υποχρεωτικό στάδιο για τη δημιουργία αυτοματοποιημένων εργαλείων της Υπολογιστικής Γλωσσολογίας (Τάντος 2015, 214).

2.5. Εργαλεία Αξιοποίησης Σωμάτων Κειμένων

Για την επεξεργασία των ΣΚ με αυτοματοποιημένο τρόπο έχουν αναπτυχθεί πολλά εργαλεία. Ένα από αυτά που είναι ευρέως διαδεδομένο και παρέχει πλήρη πρόσβαση στο διαδίκτυο δωρεάν είναι το AntConc, το οποίο θα χρησιμοποιηθεί και στο ερευνητικό κομμάτι της διπλωματικής. Πρόκειται για ένα εργαλείο όπου εισάγονται τα δεδομένα ένα - ένα με πληκτρολόγηση και στη συνέχεια το πρόγραμμα μπορεί να τα επεξεργαστεί αυτόματα. Πρόκειται επομένως για ένα ημι-αυτόματο εργαλείο επεξεργασίας δεδομένων.

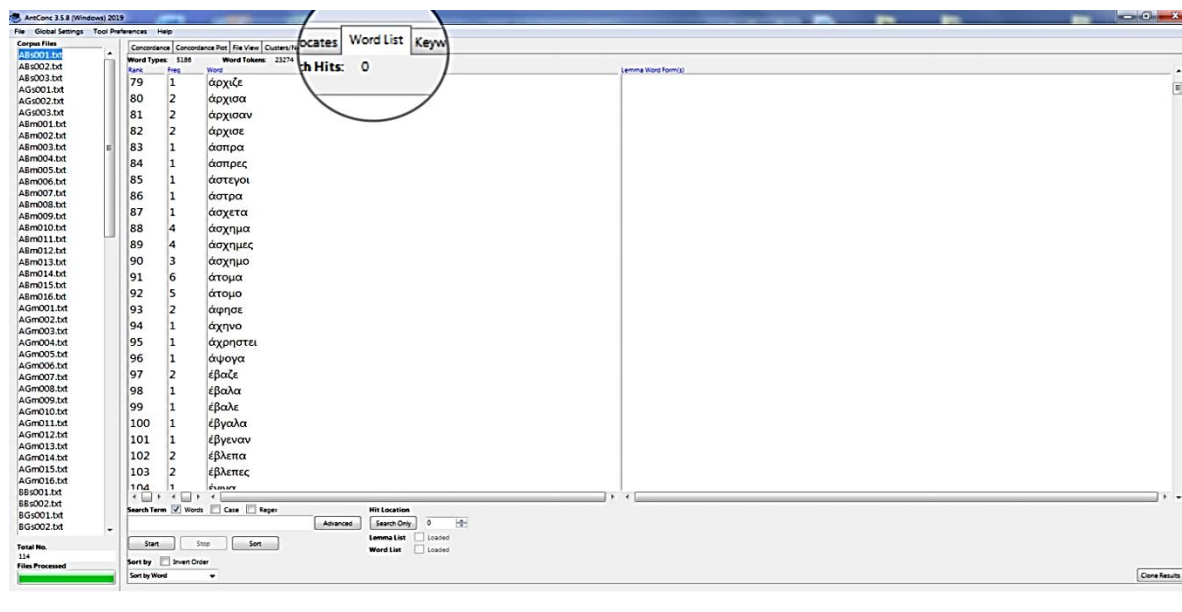
Πιο συγκεκριμένα, ένα ΣΚ μπορεί να αξιοποιηθεί κάνοντας χρήση των συμφραστικών πινάκων (concordances), των καταλόγων συχνοτήτων (word list) και των συνεμφανίσεων (Collocations-coccurrences) (Τάντος 2015, 214-219).

Συμφραστικοί πίνακες: είναι μια λίστα από αποτελέσματα όρων ή φράσεων που επιλέγουμε να αναζητήσουμε, όχι όμως αποκομμένα αλλά μαζί με τις λέξεις που συνεμφανίζονται στην πρόταση, στοιχισμένα το ένα κάτω από το άλλο τα οποία μπορούμε να επεξεργαστούμε ώστε να εξάγουμε χρήσιμες πληροφορίες.



Εικόνα 1. Στιγμιότυπο αναζήτησης του AntConc της λέξης «όταν» στο συμφραστικό πίνακα.

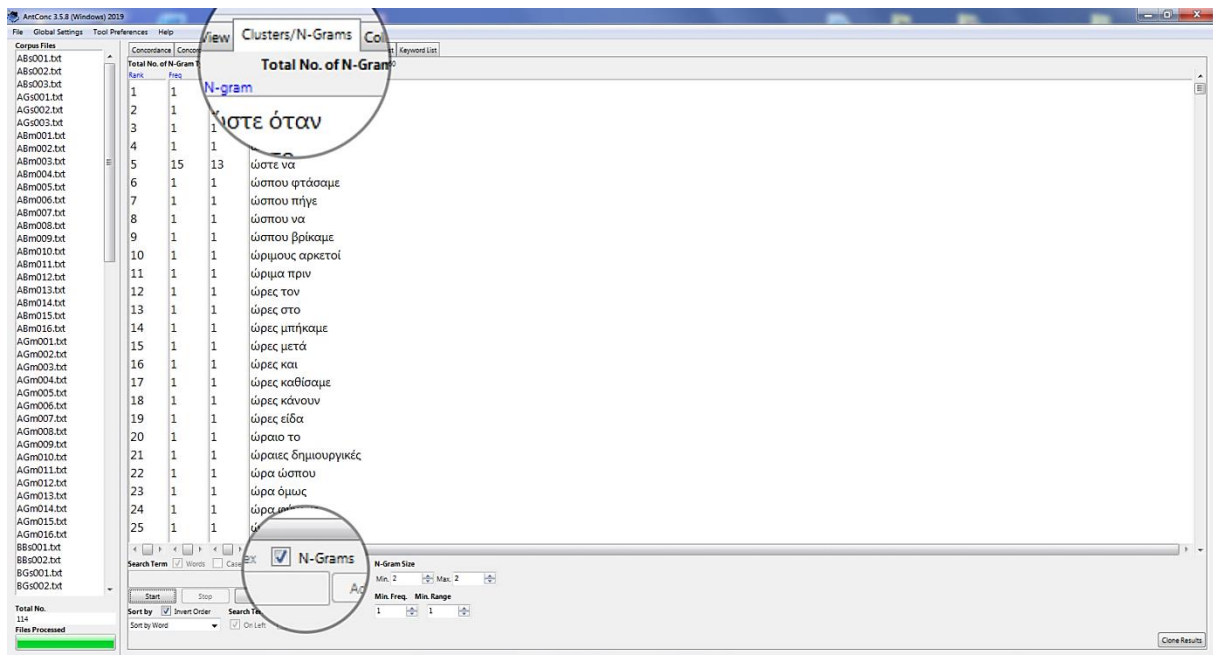
Κατάλογος συχνότητας: Η συχνότητα είναι βασικός όρος στη γλωσσολογία σωμάτων κειμένων γιατί με αυτήν συνδέεται το τι είναι πιο σημαντικό και τι δεν είναι σημαντικό στη γλώσσα. Με αναζήτηση στον κατάλογο συχνότητας μπορεί να διερευνηθεί πόσες φορές απαντά κάθε μια λέξη που βρίσκεται στο ΣΚ.



Εικόνα 2. Στιγμιότυπο αναζήτησης του AntConc στον κατάλογο συχνοτήτων

Νίγραμμα και συνεμφάνιση: Ο πιο γενικός όρος για τη συνεμφάνιση των λέξεων είναι το νίγραμμα (ngram), που εσωκλείει όλες τις περιπτώσεις συνεμφάνισεων (δίγραμμα, τρίγραμμα κτλ.). Η μελέτη των περιπτώσεων αυτών εξετάζεται από δύο διαφορετικές οπτικές σε ότι αφορά την υπάρχουσα βιβλιογραφία και τονίζει τη διαφορετική στόχευση μεταξύ υπολογιστικής και θεωρητικής γλωσσολογίας. Η πρώτη εστιάζει στην *εντασιακή* ερμηνεία (intensional) της συνεμφάνισης ως φαινομένου παραβίασης της αρχής της συνθετικότητας, ενώ η δεύτερη εστιάζει στην *εκτασιακή* ερμηνεία (extensional). Μελετάτε δηλαδή ως αποτύπωση μιας κατανομής πιθανοτήτων που μπορεί να περιέχει δύο, τρεις ή περισσότερες συνεμφάνισεις λέξεων.¹

¹ Για μια σύντομη θεωρητική αναφορά στην έννοια της συνεμφάνισης των λέξεων βλ. Bartsch (2004).



Εικόνα 3. Στιγμιότυπο αναζήτησης του AntConc στον κατάλογο νιγράμματος

2.6. Σώματα Κειμένων της Νέας Ελληνικής Γλώσσας

Τα Σώματα Κειμένων τα οποία είναι διαθέσιμα στη νέα ελληνική γλώσσα είναι αρκετά περιορισμένα ως προς το μέγεθος, κυρίως αν συγκριθούν με σώματα κειμένων της αγγλικής γλώσσας. Ωστόσο το περιεχόμενό τους είναι προσπελάσιμο στο διαδικτυακό χώρο πλήρως ή μερικώς. Τα πιο γνωστά ΣΚ στην ελληνική γλώσσα που είναι και η γλώσσα που μας ενδιαφέρει στην παρούσα διπλωματική εργασία είναι επτά. Τα τρία από αυτά είναι γενικότερης φύσεως ενώ τα άλλα τέσσερα πιο εξειδικευμένα.

Γενικά Σώματα Κειμένων:

1. [Ο Εθνικός Θησαυρός της Ελληνικής Γλώσσας \(ΕΘΕΓ\)](#) του Ινστιτούτου επεξεργασίας λόγου είναι το πολυπληθέστερο σε λέξεις σώμα κειμένων της Νέας Ελληνικής. Περιέχει πάνω από πενήντα ένα (51) εκατομμύρια λέξεις και συνεχώς εμπλουτίζεται. (Ινστιτούτο Επεξεργασίας Λόγου 2019). Τα κείμενα που εμπεριέχονται αποτελούν αντιπροσωπευτικό δείγμα της Νέας Ελληνικής και η χρονολόγησή τους ξεκινά δειλά από κείμενα του 1976 έως το 2007. (Χατζηγεωργίου κ.α. 2000, 2 · Πολίτης 2008). Η αρχειοθέτηση των κειμένων έχει βασιστεί στο πρότυπο Parole (Parole 1995) και ταξινομούνται με βάση το μέσο που δημοσιεύτηκαν (πχ. Εφημερίδες, βιβλία, περιοδικά, διάφορα), το κειμενικό είδος (άρθρο, δοκίμιο, βιογραφίες κτλ.) και το

περιεχόμενο (πχ. ιστορία, γεωγραφία, επιστήμες). Ωστόσο ως προς το κειμενικό είδος έχουν υπάρξει επιπλέον υποκατηγορίες για την καλύτερη ταξινόμηση τους. (Gavriliidou et al. 1998 στο Χατζηγεωργίου κα. 2000, 3). Στον Εθνικό θησαυρό μπορούν να αναζητηθούν λέξεις, λήμματα και γραμματικές κατηγορίες. Θεωρείται μια πηγή πλούσιων μεταδεδομένων και κειμενικών γλωσσικών χαρακτηρισμών. Ωστόσο στα μειονεκτήματα θα μπορούσε να συμπεριληφθεί η μερική πρόσβαση δεδομένων σε άτομα που δεν είναι συνδρομητές (Goutsos 2010 · Γούτσος και Φραγκάκη 2015, 53).

2. Το [Σώμα Ελληνικών Κειμένων](#) (ΣΕΚ) δημιουργήθηκε με σκοπό να βοηθήσει τη γλωσσολογική έρευνα μέσα από τη μελέτη ποικίλων κειμενικών ειδών της σύγχρονης ελληνικής γλώσσας. Είναι δημιούργημα συνεργασίας του Πανεπιστημίου Αθηνών και του Πανεπιστημίου της Κύπρου, ύστερα από τη χρηματοδότηση του προγράμματος ΠΥΘΑΓΟΡΑΣ. Το ΣΕΚ περιέχει κείμενα εκτός από γραπτά και προφορικά σε ποσοστό 10%. Το υλικό που συγκεντρώθηκε είναι από κείμενα του 1990 έως του 2010 και περιλαμβάνει τριάντα (30) εκατομμύρια λέξεις από ένα σύνολο ποικίλου κειμενικού είδους πχ. ακαδημαϊκό, λογοτεχνικό, ραδιοτηλεοπτικών συνεντεύξεων, ειδησεογραφικών κειμένων κ.α. (Goutsos 2010 · Γούτσος και Φραγκάκη 2015, 50-53).
3. Το [Gkwac](#) είναι ένα γενικό ΣΚ όπου αποκλειστική πηγή για τη δημιουργία του αποτελούν κείμενα από το διαδίκτυο. Αρχικός στόχος ήταν η αξιοποίηση του στο πεδίο των λεξικογραφικών εφαρμογών της Νέας Ελληνικής. Η πρόσβαση επιτυγχάνεται με το σύστημα αναζήτησης Sketch Engine και αποτελείται από 124 εκατομμύρια λέξεις. Πλεονέκτημα του θεωρείται η ύπαρξη γλωσσικών χαρακτηρισμών καθώς και μιας γραμματικής για την αυτοματοποιημένη συντακτική ανάλυση, επιτρέποντας με αυτόν τον τρόπο προηγμένες αναζητήσεις δεδομένων. Ωστόσο σημαντικό μειονέκτημα, εξαιτίας της σχεδίασης των κειμένων που επιλέχθηκαν κατά αποκλειστικότητα από το διαδίκτυο, είναι πως δεν μπορεί να χαρακτηριστεί από αντιπροσωπευτικότητα και ισορροπία των κειμένων που περιλαμβάνονται (Πανταζάρα 2015, 10-11).

Εξειδικευμένα Σώματα Κειμένων:

Το [Corpus Προφορικού Λόγου](#) του ιδρύματος Τριανταφυλλίδη, αποτέλεσε μια ερευνητική προσπάθεια συγκέντρωσης προφορικών δεδομένων, που πραγματοποιήθηκε στο πλαίσιο του προγράμματος «Γλωσσική Διεπίδραση και Ανάλυση Συνομιλίας» του Ινστιτούτου Νεοελληνικών σπουδών. Αρχικό μέλημα ήταν η σωστή ανάλυση της γλώσσας και η ποιοτική

επικοινωνιακή ανάλυση όπως αυτή προκύπτει μέσα από τη μελέτη της ανάλυσης της συνομιλίας. Αυτό είχε ως αποτέλεσμα να δοθεί μεγάλη προσοχή στην πιστή καταγραφή του προφορικού λόγου σε γραπτό (Pavlidou, 2012 · Ινστιτούτο Νεοελληνικών Σπουδών χ.χ.). Το corpus υπερβαίνει τις δύο (2) εκατομμύρια λέξεις μέσα από μια ποικιλία συνομιλιακού λόγου από σαράντα (40) καθημερινές συζητήσεις με φίλους και συγγενείς, 145 τηλεφωνικές ομιλίες, δεκαεπτά (17) τηλεοπτικές συζητήσεις με πολιτικά πρόσωπα καθώς και με γλωσσική διεπίδραση στη σχολική αίθουσα. Το υλικό ωστόσο που είναι προσπελάσιμο δωρεάν στο διαδίκτυο είναι συνολικού μεγέθους 671.543 λέξεων και αρκεί μια απλή εγγραφή (Ινστιτούτο Νεοελληνικών Σπουδών χ.χ.).

Το [Σώμα Νέων Ελληνικών Κειμένων](#) του Κέντρου Ελληνικής Γλώσσας (ΚΕΓ) παρέχει πρόσβαση από τη σελίδα Πύλη για την Ελληνική γλώσσα. Το σώμα απαρτίζεται από δημοσιογραφικά κείμενα των εφημερίδων ΤΑ ΝΕΑ (δύο εκατομμυρίων λέξεων) και ΜΑΚΕΔΟΝΙΑ (τριών εκατομμυρίων λέξεων) καθώς και από κείμενα των σχολικών εγχειριδίων (δύο εκατομμυρίων λέξεων). Εξαιτίας της εστίασης στο δημοσιογραφικό λόγο, επιτεύχθηκε μια λεπτομερής και λειτουργική ταξινόμηση των κειμένων ως προς την τυπολογία τους (πχ. πολιτικό, κοινωνικό, αστυνομικό, ταξιδιωτικό ρεπορτάζ κ.α.). Κάτι που το διαφοροποιεί από τα υπόλοιπα ΣΚ της Νέας Ελληνικής, είναι πως παρέχεται η δυνατότητα για μετάβαση από τον συμφραστικό πίνακα σε όλο το κείμενο και όχι σε μέρος αυτού, μέσα από μια απλή αναζήτηση λέξεων, λημμάτων ή φράσεων (Πολίτης 2008). Αυτό δίνει τη δυνατότητα να αποτελεί περιβάλλον ικανό να αξιοποιηθεί στη γλωσσική διδασκαλία, κάτι που ήταν και στους αρχικούς στόχους της δημιουργίας του (Κουτσογιάννης 2008).

Το [Σώμα Κειμένων ΙΑΤΡΟΛΕΞΗ](#), δημιουργήθηκε στη διάρκεια αυτού του ερευνητικού προγράμματος (ερευνητικό πρόγραμμα ΙΑΤΡΟΛΕΞΗ), με στόχο να δημιουργήσει την κατάλληλη υποδομή για προηγμένες εφαρμογές επεξεργασίας της φυσικής γλώσσας στο γνωστικό πεδίο της βιοϊατρικής. Για τη νέα ελληνική γλώσσα είναι το μοναδικό εξειδικευμένο σώμα, με περιεχόμενο περίπου εντεκάμιση (11,5) εκατομμύρια λέξεις. Τα κείμενα που περιέχονται είναι από ελληνικές σελίδες του διαδικτύου με ιατρικό επιστημονικό περιεχόμενο από συνέδρια, περιοδικά και εφημερίδες. Η χρονολόγηση τους είναι από το 1996-2007 και η προσπέλαση του ΣΚ επιτυγχάνεται με τη βοήθεια του εργαλείου συμφραστικών πινάκων που προσφέρει την επιλογή για αναζήτηση λημμάτων, λέξεων και γραμματικών κατηγοριών (ΙΑΤΡΟΛΕΞΗ 2006 · Tsalidis κ.α. 2007, 4-9).

Το [Ελληνικό Σώμα Κειμενικών Σχέσεων \(C58\)](#) είναι ένα χρηματοδοτούμενο έργο από το Α.Π.Θ για να καταρτιστεί και να αναλυθεί το πρώτο επισημειωμένο ΣΚ με κειμενικές σχέσεις για τα Ελληνικά (Tantos υπό έκδ.· Tantos κ.α. 2017 στο Τάντος 2015, 223-224).

Βασικός στόχος ήταν η δημιουργία ενός αυτόματου ταξινομητή κειμενικών σχέσεων μεταξύ δύο γειτνιαζόντων εκφωνημάτων. Το C58 στην πρώτη του έκδοση, αποτελείται από 58 κείμενα από δημοσιογραφικό λόγο συσταδικής δειγματοληψίας από το ΣΚ του Κέντρου Ελληνικής Γλώσσας. Με μια πρώτη ματιά φαίνεται ένα ιδιαίτερα εξειδικευμένο ΣΚ με ίσως περιορισμένο πεδίο εφαρμογής, ωστόσο μπορεί να είναι η απαρχή για μια σειρά από πρακτικές εφαρμογές ΥΓ, ή και γιατί όχι για τη βελτίωση άλλων ήδη υπάρχουσών εφαρμογών (Τάντος 2015, 223-224)

3. ΣΩΜΑΤΑ ΚΕΙΜΕΝΩΝ ΜΑΘΗΤΩΝ

3.1. Ορισμός

Η δεκαετία του '90 αποτέλεσε την απαρχή των ΜΣΚ με αντικείμενο μελέτης την αγγλική ως δεύτερη/ξένη γλώσσα (Τζιμώκας 2018, 53). Τα Μαθητικά Σώματα Κειμένων (learner corpora) είναι ένα από τα πιο εξειδικευμένα είδη ΣΚ. Πρόκειται είτε για απλές συλλογές κειμένων μαθητών/τριών σε ηλεκτρονική μορφή είτε μαθητικά κείμενα που συλλέχτηκαν και φέρουν πληροφορίες γλωσσολογικές όπως τα λάθη των μαθητών/τριών, τα μέρη του λόγου κ.α. Κατά τον Leech (1998 στο Παπαδοπούλου και Τάντος 2014, 246) τα ΜΣΚ είναι εξειδικευμένα Σώματα Κειμένων, που απαρτίζονται από ηλεκτρονικής μορφής αρχεία και αποτυπώνουν προφορικό ή/και γραπτό λόγο μαθητών/τριών μιας γλώσσας ως δεύτερης.

3.2. Μαθητικά Σώματα Κειμένων της Νέας Ελληνικής Γλώσσας ως Γ2

Στο πλαίσιο συγκεκριμένων ερευνητικών προγραμμάτων, για την ελληνική γλώσσα έχουν πραγματοποιηθεί μόλις τέσσερις προσπάθειες για τη δημιουργία και τη χρήση Σωμάτων Κειμένων Μαθητών/τριών και όλα αφορούν σώματα, που δημιουργήθηκαν από μαθητές/τριες που μάθαιναν την ελληνική ως δεύτερη γλώσσα (Γ2) (Τσιμώκας 2018, 54). Πιο συγκεκριμένα:

1) Το Εθνικό και Καποδιστριακό Πανεπιστήμιο (ΕΚΠΑ) της Αθήνας δημιούργησε το 2004, το πρώτο Ηλεκτρονικό Σώμα Κειμένων Μαθητών/τριών στο πλαίσιο του προγράμματος *Πυθαγόρας Ι*. Η προσπάθεια αφορούσε την κωδικοποίηση μιας αντιπροσωπευτικής ομάδας λαθών από ενήλικες εκπαιδευόμενους/ες, που μάθαιναν την ελληνική ως δεύτερη/ξένη γλώσσα και ανήκαν σε διαφορετικό επίπεδο γλωσσομάθειας. Το παρόν ΜΣΚ περιλαμβάνει

γύρω στις 65.000 λέξεις και πιο συγκεκριμένα 191 παραγωγές γραπτού λόγου. Βασικό του πλεονέκτημα αποτελεί η ποικιλομορφία που υπάρχει ως προς τα κειμενικά είδη. Επιπρόσθετα, στηρίζεται σε ένα πλαίσιο επισημείωσης με πολύ μεγάλη λεπτομέρεια, διευκολύνοντας την ορθότερη κατηγοριοποίηση των λαθών. Αξίζει ωστόσο να σημειωθεί ως μειονέκτημα, η συχνά κατευθυνόμενη επισημείωση των λαθών από τον επισημειωτή σε μια καθαρά υποκειμενική ερμηνεία του λάθους, που θέτει σε περιορισμό μια μελλοντική διαφορετική ερμηνεία ανάλυσης από τυχόν άλλους μελετητές-τριες/διδάσκοντες. Παρατηρείται λοιπόν έλλειψη ευελιξίας του τρόπου επισημείωσης, μειώνοντας με αυτόν τον τρόπο το εύρος χρήσης και ανάλυσης των δεδομένων και δυσκολία στον παραλληλισμό του με άλλα ΜΣΚ. Επομένως, ενώ μπορεί να έχει χρηστικότητα για ερευνητικούς σκοπούς, η διδακτική του αξιοποίηση ωστόσο συναντά δυσκολίες. (Τζιμώκας 2010, 603 · Παπαδοπούλου και Τάντος 2014, 248 · Τζιμώκας 2018, 54).

2) Το πρόγραμμα *Εκπαίδευση Αλλοδαπών και Παλινοστούτων Μαθητών* υλοποίησε το Ελληνικό Σώμα Κειμένων Μαθητών (ΕΣΚΕΙΜΑΘ) από παιδιά μεταναστών και παλινοστούτων (Παπαδοπούλου και Τάντος 2014, 248). Το υλικό συλλέχτηκε το διάστημα 2012-2013 από μαθητές/τριες όλων των βαθμίδων ηλικίας 7-19. Το μέγεθος του ΜΣΚ αποτελείται από 33.500 λέξεις ή διαφορετικά πάνω από 500 κείμενα γραπτού λόγου μαθητών/τριών. Βασική αρχή της δημιουργίας του ήταν να μπορεί να αποτελέσει ερευνητικό εργαλείο για μελλοντική χρήση από εκπαιδευτικούς/εκπαιδευόμενους-ες. Το σημαντικό πλεονέκτημα του ΕΣΚΕΙΜΑΘ σε αντιπαραβολή με τα άλλα ελληνικά ΜΣΚ είναι η ύπαρξη εξ αποστάσεως επισημείωσης. Αυτό σημαίνει πως οι επισημειώσεις και τα ψηφιοποιημένα κείμενα δεν είναι ενσωματωμένα αλλά διαχωρίζονται, προσφέροντας έτσι τη δυνατότητα πολλαπλής επισημείωσης σε ίδια δεδομένα. Επιπλέον, παρέχεται μια ευελιξία και μια επεκτασιμότητα που σε συνδυασμό με τη δυνατότητα για αποθήκευση των επισημειώσεων και των γραπτών κειμένων των μαθητών/τριών σε διαφορετικά αρχεία, βοηθάει μελλοντικούς ερευνητές/τριες για δημιουργία νέας, δικής τους επισημείωσης. (Παπαδοπούλου και Τάντος 2014, 248-249 · Τάντος κ.α. 2015, 721-723 · Τζιμώκας 2018, 54)

3) Κατά τη διάρκεια του προγράμματος *Εκπαίδευση των παιδιών της Μουσουλμανικής μειονότητας στη Θράκη* δημιουργήθηκε το αντίστοιχο ΜΣΚ. Τα κείμενα συλλέχτηκαν το διάστημα 2003-2004 και αφορούσαν καθαρά παραγωγές γραπτού λόγου Μουσουλμανοπαιδών. Το μέγεθος του εν λόγω σώματος είναι μόλις 11.300 (Τζεβελέκου κ.α. 2008).

4) Τέλος, το *Σώμα εξελικτικών Παραγωγών Μαθητών της Ελληνικής ως Γ2* (ΣΕΠΑΜΕ2), που δημιουργήθηκε από υλικό μαθητών/τριών του Διδασκαλείου της Νέας Ελληνικής ως δεύτερης/ ξένης γλώσσας του ΕΚΠΑ ύστερα από συγκέντρωση προφορικών και γραπτών κειμένων. Είναι μια προσπάθεια που ξεκίνησε τον Οκτώβριο του 2004 με σκοπό να φέρει πιο κοντά τη γλωσσολογική θεωρία των σωμάτων και την πρακτική στη δεύτερη γλώσσα. Το παρόν ΣΚ συνεχώς εμπλουτίζεται από υλικό, τα κριτήρια όμως συλλογής είναι αυστηρά. Πιο συγκεκριμένα οι συμμετέχοντες θα πρέπει να παρακολουθήσουν τα ετήσια προγράμματα Νέων Ελληνικών του ΕΚΠΑ, να επιβλέπονται για τουλάχιστον μια ακαδημαϊκή χρονιά και τέλος να κάνουν το τεστ εισαγωγής όπου βάση αποτελεσμάτων θα ενταχθούν στο αντίστοιχο επίπεδο γλωσσομάθειας (A1-C2), σύμφωνα με το Κοινό Ευρωπαϊκό Πλαίσιο Αναφοράς (2001). Μέχρι το 2016 αποτελούνταν από 650.000 λέξεις ή διαφορετικά από 1.626 παραγωγές γραπτού λόγου και 426 προφορικού. Το μέγεθος του ΣΚ σε συνδυασμό με τα αυστηρά κριτήρια σχεδιασμού βάσει των οποίων έχουν συλλεχθεί, συνιστούν βασικά πλεονεκτήματα όσον αφορά την αντιπροσωπευτικότητα των δεδομένων και την γενίκευση των αποτελεσμάτων. Αυτό προσφέρει έδαφος για νέα εργαλεία αναφοράς αλλά και για περισσότερα αξιόπιστα δεδομένα γλώσσας (Ιακώβου κ.α. 2017).

3.3. Μαθητικά Σώματα Κειμένων της Αγγλικής Γλώσσας ως Γ1

Child Language Data Exchange System (CHILDES):

Αποτελείται από περίπου είκοσι (20) εκατομμύρια λέξεις (περίπου 180 εκατομμύρια χαρακτήρες), σε είκοσι γλώσσες. Το corpus CHILDES παρέχει εργαλεία για τη μελέτη αλληλεπιδράσεων συνομιλίας. Αυτά τα εργαλεία περιλαμβάνουν μια βάση δεδομένων από μεταγραφές, προγράμματα για την ανάλυση ηλεκτρονικών μεταγραφών, μεθόδους γλωσσικής κωδικοποίησης και συστήματα για τη σύνδεση των μεταγραφών με ψηφιοποιημένο ήχο και βίντεο. Επιπλέον, περιλαμβάνει βιβλιογραφίες (Child Language Bibliographies) με ερευνητικά άρθρα σχετικά με τον τρόπο που κατακτούν τη γλώσσα τα παιδιά, και είναι διαθέσιμα σε πολλές γλώσσες (L1 και L2 Acquisition Corpora χ.χ.).

[Lancaster Corpus of Children's Project Writing \(LCPW\):](#)

Μια ψηφιοποιημένη συλλογή από εργασίες που παρήγαγαν μαθητές/τριες μεταξύ 8-11 ετών. Αποτελεί μέρος από ένα μεγάλο ερευνητικό πρόγραμμα, πιο συγκεκριμένα από μια διαχρονική μελέτη από μαθητικά γραπτά, βασισμένη στη παραγωγή παιδιών ηλικίας 8-12 ετών (L1 και L2 Acquisition Corpora χ.χ.).

Polytechnic of Wales (POW) Corpus:

Αποτελείται από 100.000 λέξεις προφορικού λόγου της Αγγλικής γλώσσας που παρήγαγαν 120 παιδιά, ηλικίας 6-12 ετών. Μια ανάλυση βασισμένη στη συστημική-λειτουργική γραμματική του Halliday. Διανεμήθηκε από το αρχείο Oxford Text Archive και το ICAME. (L1 και L2 Acquisition Corpora χ.χ.).

3.4. Μαθητικά σώματα κειμένων της Αγγλικής Γλώσσας ως Γ2

Η γλώσσα που παράγεται από μη φυσικούς ομιλητές/τριες/συγγραφείς ως ξένη/ Γ2 γλώσσα απαιτεί έρευνα. Παρακάτω παρατίθενται εν συντομία τα πιο σημαντικά Learner Corpora που δημιουργήθηκαν για την αγγλική γλώσσα ως Γ2 από μαθητευόμενους/ες.

[International Corpus of Learners English \(ICLE vers.2\):](#)

Λειτουργεί από το Μάιο του 2009, περιλαμβάνονται πάνω από 3,7 εκατομμύρια λέξεις από γραπτό λόγο προχωρημένου επιπέδου φοιτητών/τριών της αγγλικής (ως ξένης και όχι ως Γ2 γλώσσας) έχοντας δεκαέξι (16) διαφορετικές μητρικές γλώσσες. Υπάρχουν δύο τύποι συγγραφής δοκιμίου: 1) Δοκίμια (αφηγήσεις) χρησιμοποιώντας εργαλεία αναφοράς γλωσσών (λεξικά γραμματικές κα.), αλλά είναι εξ ολοκλήρου παραγωγή των μαθητών/τριών, χωρίς καμία βοήθεια φυσικού ομιλητή, 2) γραπτά εξετάσεων στο μάθημα της λογοτεχνίας (όχι περισσότερα από 25% του κάθε εθνικού σώματος). Κάθε δοκίμιο αποτελείται από 500- 1000 λέξεις. Τον Μάιο του 2009, υπήρχαν 5.554 δοκίμια και 531 λογοτεχνικά ή άλλου είδους (L1 και L2 Acquisition Corpora χ.χ.).

[Louvain International Database of Spoken English Interlanguage\(LINDSEI\):](#)

Δημιουργήθηκε το 1995 και αποτελεί ένα corpus από προφορικές συλλογές μαθητευόμενων στην αγγλική γλώσσα, έχοντας ως υπόβαθρο 11 διαφορετικές γλώσσες (Βουλγάρικα,

Κινέζικα, Ολλανδικά, Γαλλικά, Γερμανικά, Ελληνικά, Ιταλικά, Γιαπωνέζικα, Ισπανικά, Πολωνικά και Σουηδικά μέχρι σήμερα). Υπάρχουν δύο τύποι ομιλίας : ανεπίσημος λόγος σε συνεντεύξεις (ανοιχτή συζήτηση σε ένα δοθέν θέμα) και συζήτηση με εικόνες (βασισμένη σε σταθερές εικόνες). Υπάρχει ένα παρόμοιο corpus προφορικού λόγου από φυσικούς ομιλητές της αγγλικής που ονομάζεται LOCNEC (Louvain Corpus Native English Conversation), (L1 και L2 Acquisition Corpora χ.χ.).

International Corpus Network of Asian Learners of English (ICNALE):

Περιλαμβάνει πάνω από 10.000 μονολόγους, διαλόγους και δοκίμια που παράγονται από φοιτητές/τριες σε δέκα (10) χώρες/περιοχές σε όλη την Ασία (Κίνα, Hong Kong, Ινδονησία, Ιαπωνία, Κορέα, Πακιστάν, Φιλιππίνες, Σιγκαπούρη/Μαλαισία, Ταϊβάν και Ταϊλάνδη), καθώς και από τους/τις φυσικούς ομιλητές/τριες της αγγλικής γλώσσας. Σήμερα περιλαμβάνει τέσσερις ενότητες: μονολόγους, διαλόγους, γραπτά δοκίμια και επεξεργασμένα δοκίμια (L1 και L2 Acquisition Corpora χ.χ.).

Chinese Learner English Corpus (CLEC):

Αποτελείται από ένα (1) εκατομμύριο λέξεις από αγγλικές συνθέσεις που συλλέχτηκαν από πέντε διαφορετικά επίπεδα Κινέζων εκπαιδευόμενων της Αγγλικής γλώσσας με επισημειωτή, σύμφωνα με ένα σχήμα επισημείωσης 61 τύπων λαθών (εξαιρούνται τα υφολογικά λάθη τα οποία είναι δύσκολο να επισημανθούν αντικειμενικά και με συνέπεια). Αποτελείται από ένα βιβλίο και ένα CD-ROM. Το βιβλίο έχει μια παρουσίαση (στην Κινεζική γλώσσα), όπου δίνει μια περιγραφή του σχεδιασμού του σώματος, της μεθοδολογίας που χρησιμοποιείται στη στατιστική ανάλυση του σώματος και των σημαντικότερων ευρημάτων, συν μια αλφαβητική λίστα, μια λίστα λυμάτων, έναν πίνακα κατανομής λέξεων συχνότητας, έναν συνοπτικό πίνακα σφαλμάτων και έναν κατάλογο ορθογραφικών λαθών. Το CD-ROM αποτελείται από το σώμα επισημείωσης του λάθους με ένα απλό ευρετήριο και όλες τις λίστες και τους πίνακες του βιβλίου. (μη διαθέσιμο link), (L1 και L2 Acquisition Corpora χ.χ.).

The Spoken & Written English Corpus of Chinese Learners (SWECCL):

Αναπτύχθηκε υπό την ηγεσία του Qiufang Wen, του Lifei Wang, του Maocheng Liang και του Xiaojin Yan, και δημοσιεύθηκε σε CD από το Τμήμα Διδασκαλίας και Έρευνας Ξένων Γλωσσών (L1 και L2 Acquisition Corpora χ.χ.).

Spoken Learner Corpus (SLC):

Το Πρόγραμμα Spoken Learner Corpus (SLC) είναι μια συνεργασία μεταξύ του Trinity και του Κέντρου Προσεγγίσεων Κορμού στις Κοινωνικές Επιστήμες (CASS) στο Πανεπιστήμιο του Lancaster.

Στόχος του έργου είναι η δημιουργία ενός μεγάλου ΣΜΚ από προφορικά κείμενα, που θα χρησιμοποιηθούν σε ένα ευρύ φάσμα ερευνών, συμπεριλαμβανομένου του επιπέδου κατάκτησης της Γ2 γλώσσας, των γλωσσικών τεστ, καθώς και της παιδαγωγικής και υλικής ανάπτυξης στην Γ2 γλώσσα.

Το corpus, το οποίο είναι επίσης γνωστό ως *Trinity Lancaster Corpus*, περιλαμβάνει σήμερα 3,5 εκατομμύρια λέξεις και έχει δημιουργηθεί από τις ηχογραφήσεις στις εξετάσεων των αγγλικών της Trinity (GESE) σε μια κλίμακα βαθμών από το B1-C2 σύμφωνα με το Κοινό Ευρωπαϊκό Πλαίσιο Αναφοράς. Αντιπροσωπεύει τη γλώσσα που χρησιμοποιείται σε διάφορα επικοινωνιακά περιβάλλοντα στην πραγματική ζωή πέραν των τεστ και καλύπτει πολλά διαφορετικά γλωσσικά υπόβαθρα (L1 και L2 Acquisition Corpora χ.χ.).

Lancaster Corpus of Academic Written English (LANCAWE):

Ακαδημαϊκά γραπτά δείγματα από μη φυσικούς ομιλητές/τριες που μελετούν την Αγγλική γλώσσα μέσα από έναν προεπιλεγμένο και επισημειωμένο κύκλο μαθημάτων. Υπάρχει επίσης ένα μικρό υποσώμα από φυσικούς ομιλητές/τριες που μπορεί να χρησιμοποιηθεί για σύγκριση. Ορισμένα υποσώματα είναι οργανωμένα σύμφωνα με το θέμα της παραγωγής λόγου και, τα γραπτά των ομιλητών/ριών με Γ1, τις συνθήκες γραφής και τον χρόνο παραγωγής του κειμένου. Περιέχει περισσότερα από ένα κείμενα από κάθε μαθητή/τρια και αυτά περιλαμβάνουν παρόμοιες παραγωγές κειμένων που γράφει ο ίδιος/α μαθητής/τρια σε διαφορετικά χρονικά επίπεδα (π.χ. πριν, κατά τη διάρκεια και μετά τον κύκλο μαθημάτων), καθώς και σε διαφορετικά κειμενικά είδη (π.χ. περιγραφικά, διαμαρτυρίας, κλπ.) που γράφει ο ίδιος/α μαθητής/τρια στον ίδιο ή σε διαφορετικό χρόνο. Ένα εκτενές υποσώμα του LANCAWE ονομάζεται *Hinestroza-Kim Corpus (HKC)* (μη διαθέσιμο link), (L1 και L2 Acquisition Corpora χ.χ.).

Montclair Electronic Language Learners' Database (MELD):

Σώμα κειμένου μαθητών/τριών της αγγλικής ως Γ2 γλώσσας, με γραπτά κείμενα από όλα τα επίπεδα μαθητευόμενων στη Βόρεια Αμερική. Είναι ελεύθερα διαθέσιμο, κείμενα με περιορισμένο χρόνο γραφής και μη , που αρχειοθετείται έτσι ώστε η πρόοδος να μπορεί να

παρακολουθηθεί με την πάροδο του χρόνου. Επίσης, συλλέγονται δημογραφικά στοιχεία για κάθε μαθητή/τρια, συμπεριλαμβανομένης της ηλικίας, του φύλου, του ιστορικού της Γ1 και της προηγούμενης τους εμπειρίας στα αγγλικά. Τα δοκίμια επισημειώνονται συνεχώς για λάθη γραμματικής και ακαδημαϊκής γραφής, όπως αυτά καθορίζονται από μια ομάδα σχολιαστών. Η τρέχουσα βάση δεδομένων (Μάιος, 2009) αποτελείται από 44.477 λέξεις επισημειωμένου κειμένου και άλλες 53.826 λέξεις κειμένου. Επιτρέπει διάφορες αναλύσεις της γραφής των σπουδαστών, από την αξιολόγηση της προόδου με την πάροδο του χρόνου στη σχέση με το παραγόμενο λάθος και του υποβάθρου της Γ1. Τα λάθη σχολιάζονται ανεξάρτητα, από δύο εκπαιδευμένους/ες σχολιαστές/στριες χωρίς αναφορά σε προκαθορισμένη λίστα τύπων λάθους. (μη διαθέσιμο link), (Fitzpatrick και Seegmiller 2004, 3 · L1 και L2 Acquisition Corpora χ.χ).

[Vienna-Oxford International Corpus of English \(VOICE\):](#)

Ένα σώμα της αγγλικής γλώσσας ως Lingua Franca (δηλαδή, τα αγγλικά ως το μέσο επικοινωνίας που θεωρείται ως το πιο βολικό από ομιλητές/τριες από διαφορετικές Γ1 γλώσσες). Η έμφαση δίνεται στην επικοινωνία πρόσωπο με πρόσωπο των ομιλητών/τριών, μέσα από μια ευρεία γκάμα γλωσσών ως Γ1, των οποίων η πρωτοβάθμια και δευτεροβάθμια εκπαίδευση και η κοινωνικοποίηση τους δεν πραγματοποιήθηκε στην αγγλική γλώσσα. Οι ομιλίες περιλαμβάνουν ιδιωτικούς και δημόσιους διαλόγους, ιδιωτικές και δημόσιες ομαδικές συζητήσεις και περιστασιακές συνομιλίες, και συνεντεύξεις ένας προς έναν. Είναι διαθέσιμη η αναζήτηση στο VOICE 1, ωστόσο απαιτείται εγγραφή (L1 και L2 Acquisition Corpora χ.χ.).

Hong Kong University of Science & Technology (HKUST) Corpus:

Το μεγαλύτερο ΣΚ Κινέζων (Καντονέζων) μαθητευόμενων της αγγλικής γλώσσας. Αποτελείται από τριάντα (30) εκατομμύρια. λέξεις, με γραμματικές επισημειώσεις και επισημειώσεις χαρακτηριστικών γνωρισμάτων (A List of Corpora of Learner English χ.χ. · L1 και L2 Acquisition Corpora χ.χ.)

[SILS Learner Corpus of English \(Waseda University\):](#)

Δοκίμια από τους/τις φοιτητές/τριες της SILS, της Σχολής Διεθνών Φιλελεύθερων Σπουδών στο Πανεπιστήμιο Waseda, της Ιαπωνίας. Υπάρχει μεγάλη ποικιλία ως προς την μητρική γλώσσα, με την πλειονότητα των φοιτητών/τριών να μιλά την Ιαπωνική. Ως Μαθητικό

Σώμα μπορεί να χρησιμοποιηθεί για να εξετάσει τις επιδράσεις της μητρικής (Γ1) γλώσσας και του εκπαιδευτικού υπόβαθρου στις δεξιότητες γραφής στην αγγλική γλώσσα. Συγκεντρώνει πολλά δοκίμια από κάθε μαθητή/τρια (παρέχοντας διαχρονικότητα), τόσο για να εξετάσει το πρώτο ζητούμενο όσο και το δεύτερο, με δυνατότητα των εκπαιδευτικών να κάνουν τις παρατηρήσεις τους στα γραπτά (SILS Learner Corpus of English 2007).

[Thai English Learner Corpus \(TLEC\):](#)

Γραπτό σώμα κειμένου με 1,3 εκατομμύριο λέξεις (μέχρι 23/5/2002), με επισημείωση για τα μέρη του λόγου και λήμματα. Περιέχει δείγματα γραφής από φοιτητές/τριες του Πανεπιστημίου της Ταϊλάνδης από το 1997 και συνεχίζει να αυξάνεται. Αποτελείται από 700.000 λέξεις γραπτού λόγου στην αγγλική γλώσσα από τις εισαγωγικές εξετάσεις στο πανεπιστήμιο στο Ινστιτούτο Αγγλικής Γλώσσας (IELE, University of Assumption της Ταϊλάνδης) και 600.000 λέξεις από δοκίμια που γράφτηκαν από τους/τις μαθητές/τριες του Thai EFL. Παρέχεται δυνατότητα δωρεάν αναζήτησης στο διαδίκτυο, αλλά περιορίζεται σε εκατό (100) γραμμές ευρετηρίου. Για πλήρη πρόσβαση, πρέπει να υπάρχει επικοινωνία με τους ιδιοκτήτες (L1 και L2 Acquisition Corpora χ.χ.).

Tswana Learner English Corpus (TLEC):

Η δημιουργία του στηρίχτηκε στο ICLE. Αποτελείται από ένα σύνολο 200.000 λέξεων από δοκίμια προχωρημένου επιπέδου μαθητών/τριών της Αγγλικής από ιδρύματα της Νότιας Αφρικής. (μη διαθέσιμο link), (L1 και L2 Acquisition Corpora χ.χ.)

[Longman Learners' Corpus:](#)

Δεν είναι ελεύθερα προσβάσιμο (εκτός και αν γίνει συμφωνία με τους εκδότες). Οι σπουδαστές/τριες και οι καθηγητές/τριες απ' όλο τον κόσμο έστειλαν δοκίμια και γραπτά εξετάσεων για να βοηθήσουν στη δημιουργία του Longman Learners' Corpus. Μιας μηχανογραφημένης βάσης δεδομένων δέκα (10) εκατομμύρια λέξεων που αποτελείται εξ ολοκλήρου από γραπτά μαθητών/τριών στην αγγλική γλώσσα. Κάθε εθνικότητα, κάθε επίπεδο γλώσσας εκπροσωπείται στο σώμα κειμένου παρέχοντας μια μοναδική εικόνα για τους μαθητές/τριες για γνώση της Αγγλικής. (μη διαθέσιμο link). (L1 και L2 Acquisition Corpora χ.χ.)

Cambridge Learner Corpus (CLC):

Δεν είναι γενικά προσβάσιμο. Πρόκειται για μια μεγάλη συλλογή παραδειγμάτων από γραπτά μαθητών/τριών στην Αγγλική γλώσσα απ' όλο τον κόσμο. Πάνω από δεκαπέντε (15) εκατομμύρια λέξεις και επεκτείνονται συνεχώς. Αποτελεί τμήμα του Cambridge International Corpus (CIC). Προέρχεται από ανώνυμα γραπτά εξετάσεων από μαθητές/τριες που έδωσαν εξετάσεις αγγλικών για το πτυχίο Cambridge ESOL απ' όλο τον κόσμο. Κάθε γραπτό είναι κωδικοποιημένο με πληροφορίες σχετικά με την πρώτη γλώσσα του/της μαθητή/τριας, την εθνικότητα, το επίπεδο αγγλικής γλώσσας, την ηλικία κλπ. Αυτή τη στιγμή μπορεί να χρησιμοποιηθεί μόνο από συγγραφείς και μελετητές/τριες γραπτών που εργάζονται για το Cambridge University Press καθώς και από μέλη του προσωπικού του Cambridge ESOL (L1 και L2 Acquisition Corpora χ.χ.)

4. ΒΑΣΙΚΕΣ ΕΦΑΡΜΟΓΕΣ ΤΩΝ ΣΩΜΑΤΩΝ ΚΕΙΜΕΝΩΝ ΣΤΗ ΓΛΩΣΣΙΚΗ ΔΙΔΑΣΚΑΛΙΑ

4.1. Η Συμβολή της Γλωσσολογίας Σωμάτων Κειμένων στη Γλωσσική Διδασκαλία

Η ανέλιξη των ΣΚ ήταν ταχύτερη τις τελευταίες δεκαετίες. Αυτό έχει άμεση σχέση με την ταυτόχρονη εξάπλωση και διάδοση του ηλεκτρονικού υπολογιστή, που είναι και το βασικό εργαλείο χρήσης των ΣΚ. Η συνεισφορά τους στο πεδίο της έρευνας ήταν πολύ σημαντική αυτά τα χρόνια, αφού απαρτίζει τη βάση για να μπορέσει να αναλυθεί και να χαρτογραφηθεί η γλώσσα. Επιπλέον, προσέφερε σε ερευνητές/τριες την ευκαιρία να εξετάσουν σε βάθος τις διαφορετικές γλωσσικές ποικιλίες. Τα ΣΚ απαρτίζουν υλικό αυθόρμητου λόγου που σε συνδυασμό με το φυσικό τρόπο συλλογής των δεδομένων και όχι με τρόπο που προκύπτει μέσα από ερευνητικές δράσεις, αποτελεί βοήθεια για τους/τις ερευνητές/τριες ώστε να προβούν σε συμπεράσματα σχετικά με αυτό που λέγεται και αυτό που ισχύει σε μια γλώσσα (Γούτσος και Φραγκάκη 2015).

Το ενδιαφέρον της διεθνούς βιβλιογραφίας, είναι στραμμένο κυρίως στις δυνατότητες αξιοποίησης τους κατά την εκμάθηση και διδασκαλία κυρίως της αγγλικής ως δεύτερης/ξένης γλώσσας, χωρίς αυτό να περιθωριοποιεί τις άλλες ξένες γλώσσες (Giagkou και

Antoniou- Kritikou 2008, 387 · Αντωνίου-Κρητικού κ.α. 2010, 102). Το πεδίο εφαρμογής τους επικεντρώνεται i) στη δημιουργία διδακτικών εγχειριδίων, ii) στη σύνταξη γραμματικής και λεξικών ως εγχειρίδια αναφοράς, iii) στην άμεση χρήση μαθητών/τριών των ΣΚ εντός τάξης με σκοπό την πραγμάτωση δραστηριοτήτων, iv) την προετοιμασία υλικού για διδασκαλία από τον ίδιο τον/την εκπαιδευτικό για εφαρμογή στην τάξη. Η έρευνα εστιάζει κυρίως στη χρήση του ΣΚ που στόχο έχει να αποκτήσει ενημερότητα ο μαθητής/τρια ως μη φυσικός/ή ομιλητής/τρια της γλώσσας για τις μορφές της που θεωρούνται αυτονόητες μόνο σε έναν φυσικό ομιλητή. Υποστηρίζεται επίσης, πως η συστηματική μελέτη Σωμάτων Κειμένου που παρήχθησαν από φυσικούς ομιλητές/τριες, προσφέρει σημαντική βοήθεια σε άτομα που θέλουν να μάθουν τη γλώσσα στόχο, προσφέροντας τους εξοικείωση μέσα από διάφορες επικοινωνιακές περιστάσεις (Carter 1998 · Mair 2002 · O'Keefe McCarthy και Carter 2010).

Συμπερασματικά, μπορούμε να πούμε πως τα ΣΚ άσκησαν τεράστια επιρροή στο χώρο της εφαρμοσμένης γλωσσολογίας καθώς μετέβαλαν τον τρόπο αντίληψης και κατανόησης της γλώσσας μέσα από το πρίσμα των διαφορετικών πεδίων μελέτης. Επιπλέον, προσέφεραν μια νέα εμπειρική οπτική κατά τη γλωσσική μελέτη, κάτι μη εφικτό παλαιότερα, είτε λόγω πλήρους αδυναμίας είτε μερικής, αφού επέτρεψαν να περιθωριοποιηθεί η διαίσθηση, να εμπλουτιστούν οι ερμηνείες, να ενισχυθούν οι ισχυρισμοί και γενικά οι ερευνητές/τριες να μιλούν για τη γλώσσα με πιο μεγάλη αυτοπεποίθηση.

4.2. Η αρχή της γλωσσοδιδασκτικής αξιοποίησης των Σωμάτων Κειμένου

Η έκδοση του πρώτου λεξικού Collins COBUILD English Language Dictionary (1987), που δημιουργήθηκε στην αγγλική γλώσσα και στήριξε τις ερμηνείες του σε Ηλεκτρονικό Σώμα Κειμένων, αποτέλεσε μεγάλη πρωτοπορία της εποχής. Το παρόν λεξικό απευθύνθηκε σε άτομα που διδάσκονται την αγγλική γλώσσα ως δεύτερη ή ξένη. Ο Tim Johns, συνεργάτης του Sinclair στη δημιουργία του COBUILD πρότζεκτ, ήταν αυτός που εστίασε στην αξιοποίηση των corpora, τονίζοντας πως τα δεδομένα από αυθεντικά περιβάλλοντα που χρησιμοποιούνται από τους λεξικογράφους και τους συγγραφείς για διδακτικό υλικό, δε χρειάζεται να αποκρύπτεται από τους διδασκόμενους, αντιθέτως καλό είναι να έρχονται σε επαφή με αυτά και να αποκτούν το ρόλο του/της ερευνητή/τριας (Johns 2002 στο Χαλισιάνη 2009, 820). Ύστερα από τη διατύπωση του Johns για τη χρήση συμφραστικών πινάκων στη διδασκαλία της γλώσσας δεν ήταν λίγοι αυτοί που αντιλήφθηκαν τα ΣΚ ως ένα μέσο που

μπορεί να φέρει ριζικές αλλαγές στο πεδίο της γλωσσικής εκπαίδευσης μιλώντας μάλιστα για «επανάσταση των Σωμάτων Κειμένων» (Rundell και Stock 1992 · Rundel 2008· McCarthy 2008). Ο Johns εισήγαγε τον όρο «data driven learning» για την περιγραφή της γνώσης μιας γλώσσας που αποκτάται μέσα από την επεξεργασία ενός σώματος κειμένων. Ένας όρος που αποτελεί θεμελιώδη παραδοχή για όσους υποδεικνύουν προτάσεις για αξιοποίηση των ΣΚ στη διδασκαλία. Δημοσιεύσεις με μεγαλύτερο ενδιαφέρον για τη γλωσσική διδασκαλία μέσω ΗΣΚ κάνουν την εμφάνιση τους κυρίως από το 1997 και εξής² (Johns 2002 στο Χαλισιάνη 2009, 820).

4.3. Η αξιοποίηση των Σωμάτων Κειμένων στη Γλωσσική Διδασκαλία

Οι δυνατότητες για αξιοποίηση των Ηλεκτρονικών Σωμάτων Κειμένων στη διδασκαλία της γλώσσας αρχίζουν να γίνονται αποδεκτές από την επιστημονική κοινότητα των γλωσσολόγων, των θεωρητικών της εκπαίδευσης καθώς και των εκπαιδευτικών γενικότερα, γύρω στα τέλη της δεκαετίας του 1980, όπως προαναφέρθηκε. Καθοριστικός παράγοντας σε αυτό υπήρξε η παραδοχή της σπουδαιότητας τους στη γλωσσολογική έρευνα για την περιγραφή μιας γλώσσας αλλά και σε εφαρμοσμένους τομείς, όπως λόγου χάρη τη λεξικογραφία. Άρχισε να γίνεται κατανοητό πως το γλωσσικό υλικό από αυθεντικό περιβάλλον που περιέχουν, συνδυαστικά με την εύκολη τεχνολογικά προσπέλασή τους, αναδεικνύουν τα ΣΚ αξιοποιήσιμα, τόσο στη μελέτη και την περιγραφή όσο και στη διδασκαλία της γλώσσας. Τα ΣΚ που εφαρμόστηκαν στη γλωσσική διδασκαλία εντάχθηκαν στον κλάδο της εφαρμοσμένης γλωσσολογίας («Linguistics applied»), (Widdowson 2000 στο Χαλισιάνη 2009, 819). Οι αναζητήσεις στη θεωρητική έρευνα για την περιγραφή των γλωσσών εξέφρασαν την τάση να εμπνέουν τη δραστηριότητα σε εφαρμοσμένους κλάδους όπως είναι ο κλάδος της εκπαιδευτικής γλωσσολογίας (Braun 2005). Αυτό εξηγεί γιατί τα μοντέλα διδακτικής αξιοποίησης των Σωμάτων Κειμένων που προτείνονται, συγκλείουν σε μεγάλο ποσοστό με τα ενδιαφέροντα που καλλιεργούνται κατά τη γλωσσολογική μελέτη των ΣΚ και εκφράζουν συχνά μοντέλα γλωσσικής περιγραφής του χώρου αυτού (Kaltenböck και Mehlmauer-Larcher 2005, 77-78).

² Εκείνη τη χρονιά εκδόθηκαν και δύο συλλογικοί τόμοι που συμπεριέλαβαν πολύ σημαντικά άρθρα. (Wichmann, Fligelstone, McEnery και Knowles 2007 · Lewandowska-Tomaszczyk και Melia 1997).

Στο εκπαιδευτικό σύστημα της Ελλάδας, τα γλωσσικά εγχειρίδια στο Γυμνάσιο (σειρά 2006) σημειώνουν στροφή προς το κείμενο. Αυτό επιτυγχάνεται μέσα από την εστίαση α) στην περιγραφή β) την αφήγηση και γ) την επιχειρηματολογία, αλλά και μέσα από την αυξητική τάση των κειμένων που εμπεριέχονται σε κάθε ενότητα. Στόχος είναι να αναβαθμιστεί ο ρόλος του κειμένου ώστε να εκπληρωθούν οι διδακτικοί στόχοι κάθε διδακτικής ενότητας. Παρά το εγχείρημα αυτό, η οργάνωση της ύλης συνεχίζει να εστιάζει σε λεξικογραμματικά φαινόμενα και στην κατάτμηση της ύλης στους τρεις βασικούς άξονες γλωσσικής ανάλυσης: τη γραμματική (μορφολογία, γραμματικές κατηγορίες), τη σύνταξη και το λεξιλόγιο (Χατζησαββίδης και Χατζησαββίδου 2009).

Γραμματική: Η διδακτική αξιοποίηση των ΣΚ επικεντρώνεται συχνά στη δομή της πρότασης ή αλλιώς στη γραμματική της αλλά και σε δομές που παρουσιάζουν υψηλά ποσοστά εμφάνισης, τουλάχιστον για κάποιες ποικιλίες του λόγου. Αυτό που εγείρει το ενδιαφέρον με τη γραμματική, είναι πως δεν πρόκειται για μια κλασική προσέγγιση «εστίασης στον τύπο»³, αλλά προσπαθεί να προβάλλει τη διαφοροποίηση των προτασιακών δομών αναλόγως του κειμενικού είδους και της λειτουργικής ποικιλίας του λόγου. Εστιάζει θα λέγαμε διαφορετικά, στη συσχέτιση των γραμματικών δομών με το λόγο ως ένα γεγονός ολοκληρωμένο με το πλαίσιο παραγωγής του. Η εστίαση δε γίνεται στις σωστές ή όχι γραμματικές δομές, αλλά στον επιτονισμό διαφορετικής/εναλλακτικής γραμματικής δομής που παρουσιάζεται σε ένα ΣΚ. (Congrad 2000 στο Χαλυσιάνη 2008, 822 · Χαλυσιάνη 2010).

Λεξιλόγιο: Ένα ΣΚ έχει άμεση σχέση με τις λέξεις και κατ' επέκταση με το λεξιλόγιο αφού η αναζήτηση σε αυτό ως επί το πλείστον παρέχει πληροφορίες για την εμφάνιση, τη λειτουργία και τη χρήση των λέξεων στα διάφορα συμφραστικά περιβάλλοντα. Οι πίνακες συμφραστικού περιεχομένου (concordances), αποτελούν ιδανικό εργαλείο για τη βαθύτερη κατανόηση και κατά συνέπεια για τη διδασκαλία της σημασίας λέξεων μέσα από το περιβάλλον χρήσης τους (keyword in context). Βασική όψη της διδασκαλίας λεξιλογίου στο ΣΚ αποτελεί η μελέτη των λιγότερο συζητημένων και προφανών μορφών της ερμηνείας και του τρόπου χρήσης τους, όπως αυτές παρουσιάζονται στις λεκτικές ακολουθίες και συμπεριφράσεις (collocations) που συναντώνται. Πιο σπάνια η προσοχή εστιάζεται στη σημασία των λέξεων στο περιεχόμενο τους (context) και ως τέτοιο υπολογίζεται το είδος του κειμένου (genre) είτε η λειτουργική ποικιλία του λόγου που αντιπροσωπεύεται από αυτό το κείμενο (register). Σε αυτό που συγκλίνουν οι απόψεις των μελετητών/τριών για την

³ Εστίαση στον τύπο: στοχεύει να στρέψει την προσοχή του/της μαθητή/τριας σε συγκεκριμένες μορφές κατά τη διάρκεια μιας δραστηριότητας επικοινωνιακού χαρακτήρα (Long 1991).

αξιοποίηση των ΣΚ στη γραμματική και το λεξιλόγιο είναι η ρευστότητα των ορίων ανάμεσα σε αυτά τα δύο: τα δεδομένα που εμφανίζουν οι συμφραστικοί πίνακες παρουσιάζουν μορφές της γλωσσικής χρήσης που καθιστούν τα όρια ανάλυσης του λεξιλογίου και της γραμματικής δυσδιάκριτα ή ρευστά (Χαλσιάνη 2008, 822 · 2010 · 2011, 115).

Σύνταξη: Μέσα από τα ΣΚ μπορούν να μελετηθούν και να κατανοηθούν όλα σχεδόν τα συντακτικά φαινόμενα με την προϋπόθεση να γίνει χρήση κατάλληλων όρων αναζήτησης ώστε να εντοπιστούν τα επιθυμητά παραδείγματα. Ο/Η εκπαιδευτικός που θα προτιμήσει στην τάξη να κάνει χρήση συμφραστικών πινάκων, μπορεί να ακολουθήσει δύο μονοπάτια. Το πρώτο είναι να θεωρήσει τα αποσπάσματα από το ΣΚ δεδομένα και να ζητήσει από τους/τις μαθητές/τριες να τα επεξεργαστούν, ώστε να κατανοήσουν τις δομές της πρότασης και τις συντακτικές λειτουργίες εμποδώνοντας με αυτόν τρόπο τα συντακτικά φαινόμενα.. Το δεύτερο είναι να προτρέψει τους/τις μαθητές/τριες να διαχειριστούν τα συντακτικά φαινόμενα όχι ως καθολικά χαρακτηριστικά της μητρικής τους γλώσσας, αλλά ως ευέλικτες δομές της γλώσσας που μπορούν να προσαρμοστούν στα νοήματα που επιδιώκεται με τη χρήση της να δηλωθούν. Ακόμα και αν το ΣΚ δεν έχει ιδιαίτερα μεγάλο όγκο, τα παραδείγματα που μπορούν να μελετηθούν είναι αρκετά ώστε να βρεθεί υλικό κατάλληλο για το κοινό στο οποίο απευθύνεται, την περίσταση, τις συμβάσεις του κειμενικού είδους, το ύφος που θέλει να προσδοθεί στο εκφώνημα κ.α (Χαλσιάνη 2010 · 2011, 111).

Εστίαση στη γλωσσική χρήση: Η αξιοποίηση υλικού από τα ΣΚ για το μάθημα της γλωσσικής διδασκαλίας, στρέφει το ενδιαφέρον από την ιδεατή μορφή της γλώσσας ως πρότυπο, στην αυθεντική χρήση της, όπως εμφανίζεται στα δείγματα του λόγου που υπάρχουν στα ΣΚ (Roemer 2004 στο Χαλσιάνη 2008, 823)

Εστίαση σε ζητήματα γλωσσικής ποικιλίας: Αναδεικνύονται επιμέρους μορφές της γλωσσικής ποικιλότητας μέσα από τα ΣΚ. Αυτό μπορεί να αφορά αποκλειστικά την ποικιλία στις γραμματικές δομές (Clear 2000 στο Χαλσιάνη 2008, 823), την ποικιλία της χρήσης και της σημασίας του ίδιου γλωσσικού στοιχείου (πχ. λέξη, φραστική δομή), όπως παρουσιάζεται στα διαφορετικού λόγου δείγματα (Aston 2000 στο Χαλσιάνη 2008, 823), είτε ακόμα και τα ιδιαίτερα λεξικογραμματικά γνωρίσματα στα διάφορα κειμενικά είδη ή τέλος στις λειτουργικές τους ποικιλίες (genres και registers).

Διδασκαλία κειμενικών ειδών: Ένα ΣΚ που είναι δομημένο σε υποσώματα με ταξινόμηση επιμέρους ποικιλιών λόγου, αφήνει ανοιχτούς ορίζοντες στη γλωσσική διδασκαλία με εστίαση, αφού επιτρέπει την λεπτομερή συσχέτιση των λεξικογραμματικών επιλογών με

κάποιο κειμενικό είδος ή γενικότερα με οποιαδήποτε ποικιλία λόγου. Στο Λύκειο που το μάθημα της Έκθεσης- Έκφρασης στρέφεται προς την κειμενοκεντρική προσέγγιση μέσα από την εστίαση σε συγκεκριμένα κειμενικά είδη (είδηση, δοκίμιο, βιογραφία κ.α), το ΣΚ της Πύλης του Κέντρου Ελληνικής Γλώσσας μπορεί να αξιοποιηθεί μέσα από την εστίαση του κειμένου στο μάθημα της γλωσσικής διδασκαλίας. Μπορεί από τη μία να λειτουργήσει ως ένας χώρος αποθήκευσης κειμένων που ανά πάσα στιγμή μπορούν να ανασυρθούν αντιπροσωπευτικά δείγματα από μια ποικιλία κειμενικών ειδών ή τύπων αναλόγως του διδακτικού στόχου. Μπορεί από την άλλη, να λειτουργήσει ως εργαλείο δημιουργίας συμφραστικών πινάκων που θα αντικατοπτρίζουν τη διαφορετική χρήση του ίδιου γλωσσικού στοιχείου σε άλλα κειμενικά είδη. Ο/Η εκπαιδευτικός μέσα από διάφορες στρατηγικές μπορεί να κατευθύνει τη διδασκαλία του. Για παράδειγμα με αφορμή μια περιγραφή ενός κειμενικού είδους (ως προς τη δόμη, το λεξιλόγιο, το επίπεδο ύφους, τις συντακτικές δομές), να προχωρήσει με τους μαθητές/τριες στο ΣΚ και να αντλήσουν παραδείγματα που ανήκουν στο συγκεκριμένο κειμενικό είδος και να διερευνήσουν τον βαθμό σύγκλισης και απόκλισης από τον κανόνα (Χαλσιάνη 2008, 822-823 · 2010 · 2011, 126-129).

Παρόλο που οι εφαρμογές των ΣΚ όπως έγινε φανερό μπορούν να είναι πολλές και ευεργετικές για τη διδασκαλία, ωστόσο μέχρι στιγμής δεν έχουν ενταχθεί συστηματικά στη διδασκαλία για τη μητρική γλώσσα αλλά ούτε και στη διδασκαλία της ελληνικής ως δεύτερης/ξένης γλώσσας. Οι φιλόλογοι της δευτεροβάθμιας εκπαίδευσης λαμβάνουν ενημέρωση για τις δυνατότητες που προσφέρουν τα ΣΚ μόνο μέσα από προγράμματα επιμόρφωσης για τη διδακτική αξιοποίηση των Τεχνολογιών της Πληροφορίας και της Επικοινωνίας (ΤΠΕ), (Χαλσιάνη 2009, 826).

4.4. Πλεονεκτήματα στη διδασκαλία με Μαθητικά Σώματα Κειμένων

Τα Σώματα Κειμένων που δημιουργούνται από φυσικούς/ες ομιλητές/τριες αποτελούν ένα αυθεντικό περιβάλλον χρήσης της γλώσσας με δυνατότητα συνεχούς πρόσβασης. Επιτρέπουν λοιπόν το περιεχόμενο της γλωσσικής διδασκαλίας να προσαρμόζεται σε ένα πραγματικό γλωσσικό περιβάλλον χρήσης και όχι σε ιδεατά κατασκευασμένα παραδείγματα, αποδεσμεύοντας με αυτό τον τρόπο τη διδασκαλία από γλωσσικούς κανόνες ή διαισθητικές προσεγγίσεις των μελετητών ή των εκπαιδευτικών μιας γλώσσας (Sinclair 1997, 32-34). Επιπλέον, η πρόσβαση σε ένα ΣΚ γίνεται με τη βοήθεια εργαλείων γλωσσικής τεχνολογίας,

όπου προσφέρουν στους συνδεδεμένους χρήστες πολλούς διαφορετικούς τρόπους για αναζήτηση πληροφορίας. Αυτό τα ορίζει ως ένα ευέλικτο εργαλείο έρευνας προσπελάσιμο σε αυθεντικό γλωσσικό υλικό, ικανό να αξιοποιηθεί με διαφορετικούς τρόπους για την εκμάθηση και διδασκαλία μιας γλώσσας. Στη διεθνή βιβλιογραφία, η χρήση των corpora εστιάζει κυρίως στη ικανότητα τους για αξιοποίηση διερευνητικών πρακτικών μάθησης καθώς και στην ενθάρρυνση της αυτονομίας του μαθητή/τριας (Kaltenböck κ.α. 2005 · Braun 2005). Τα οφέλη που προκύπτουν για τους/τις μαθητευόμενους/νες μιας τάξης που χρησιμοποιεί σώματα κειμένων είναι πρωτίστως η κατάκτηση μορφών της γλωσσικής πολυμορφίας, που παρουσιάζεται σε δείγματα αυθεντικού λόγου και ακολούθως η ευκαιρία για συστηματική μελέτη της σχέσης των λεξικογραμματικών μορφών του μέσα από τη μελέτη ποικίλων κειμενικών ειδών (genre), το επίπεδο ύφους (register) του κειμένου καθώς επίσης και το κοινωνικό - επικοινωνιακό επίπεδο λόγου που πραγματώνεται (Ο' Keefe κ.α. 2010).

4.5. Αξιοποίηση Μαθητικού Σώματος Κειμένων στη γλωσσική διδασκαλία

Ο λόγος που παράγεται από τους/τις μαθητές/ριες αποτελεί μια από τις πιο βασικές μορφές ανατροφοδότησης για τον/την εκπαιδευτικό ενώ παράλληλα επιβεβαιώνει ή διαψεύδει τις γλωσσικές τους γνώσεις. Αυτό ισχύει τόσο για τη διδασκαλία της πρώτης (Γ1), όσο και για τη διδασκαλία της δεύτερης/ξένης γλώσσας (Γ2). Χωρίς τη χρήση της τεχνολογίας ο/η εκπαιδευτικός για χρόνια είχε δυνατότητα πρόσβασης σε γραπτά κείμενα μόνο μαθητών/τριών της τάξης του. Ωστόσο με τη βοήθεια της σύγχρονης τεχνολογίας, πλέον προσφέρεται η δυνατότητα πρόσβασης των εκπαιδευτικών σε γραπτό λόγο περισσότερων μαθητών/τριών και όχι απαραίτητα δικών τους μόνο. Επιπλέον, οι ίδιοι/ες μπορούν να προωθήσουν δικό τους υλικό σε άλλους συναδέλφους ως βοήθεια. Αυτό γίνεται πράξη μέσα από τη χρήση των Μαθητικών Σωμάτων Κειμένων (ΜΣΚ), δηλαδή από τις ψηφιοποιημένες συλλογές αυθεντικών μαθητικών κειμένων, που μπορούν να αξιοποιηθούν για ερευνητικούς και διδακτικούς σκοπούς. (Granger 2002, 2003b, 538,543 · Τζιμώκας 2010, 602).

Για την αξιοποίηση των ΜΣΚ υπάρχουν δύο βασικές μέθοδοι. Η αντιπαραβολική ανάλυση και η υπολογιστική ανάλυση λαθών (Granger 2002, 2003, 2008). Η πρώτη προτείνεται για τη συγκριτική μελέτη του λόγου που παράγουν μαθητές/τριες μιας Γ2 γλώσσας με τον λόγο μαθητών/τριών που έχουν την ίδια γλώσσα ως Γ1, είτε με τον λόγο από άλλους μαθητές/τριες που μαθαίνουν την ίδια γλώσσα ως Γ2 όμως έχουν διαφορετική Γ1, ηλικία κτλ.

Αντιπαραβολική ανάλυση μπορεί να υλοποιηθεί και με σκοπό να εντοπιστούν τυχόν παρεμβολές από τη Γ2 στη Γ1 γλώσσα. Από την άλλη, η υπολογιστική ανάλυση λαθών αφορά τον εντοπισμό και την κατηγοριοποίηση των γλωσσικών «λαθών» των μαθητών/τριών, με σκοπό να αντιμετωπιστούν στη διδακτική παρέμβαση.

Από τα παραπάνω συμπεραίνεται πως τα ΜΣΚ έχουν πολλές εφαρμογές και τρόπους χρήσης (Granger, 2008 · Römer, 2008 · Τζιμώκας, 2010):

Τα ΜΣΚ είναι σημαντική πηγή πληροφόρησης για τη διαγλώσσα των μαθητών/τριών δηλαδή για τις γλωσσικές ποικιλίες που κάνουν χρήση καθώς προχωρούν στα διαφορετικά επίπεδα γλωσσομάθειας. Αυτή η γνώση για τους ερευνητές/τριες από τη μία παρέχει πληροφορίες για τη μελέτη του τρόπου και των διαδοχικών σταδίων που απαιτούνται για την κατάκτηση και εκμάθηση μιας γλώσσας, και απ' την άλλη βοηθάει στο σχεδιασμό και την αξιολόγηση της εκπαιδευτικής διαδικασίας.

Τα ΜΣΜ δίνουν τη δυνατότητα να αξιοποιηθούν όχι μόνο από ερευνητές/τριες αλλά και από τους/τις εκπαιδευτικούς ως ένα μέσο για την ανίχνευση των δυσκολιών/αναγκών των μαθητών/τριών και έπειτα να λειτουργήσουν βοηθητικά για τη διαμόρφωση των διδακτικών παρεμβάσεων των εκπαιδευτικών (Granger 1997).

Επιπλέον, υπάρχει η δυνατότητα τα ίδια ΜΣΚ να λειτουργήσουν ως διδακτικό υλικό επεξεργασίας από τους μαθητές/τριες εντός τάξης. Είτε με ηλεκτρονική πρόσβαση είτε με εκτύπωση συγκεκριμένων αποσπασμάτων ή κειμένων, ο/η εκπαιδευτικός μπορεί να εξηγεί μέσα από αυθεντικά παραδείγματα τα προς διδασκαλία γλωσσικά φαινόμενα, προλαμβάνοντας ή διορθώνοντας πιθανές αποκλίσεις των μαθητών/τριών (Leech 1997 · Römer 2008 · Farr 2010 · Vaughan 2010). Πέραν από αυτό, προτείνεται και η συγκριτική μελέτη δεδομένων από την ίδια γλώσσα ως Γ2 και Γ1. Από πιο σύγχρονες μελέτες προτείνεται η δημιουργία τοπικών ΜΣΚ από εκπαιδευτικούς και μαθητές/τριες για να χρησιμοποιούνται στην τάξη, εντός σχολείου ή ακόμη και σε περισσότερα του ενός σχολείου χωρίς να είναι απαραίτητη η προσπέλαση στο ευρύ κοινό ή σε μεγαλύτερους μαθητικούς πληθυσμούς. Σε μία βάση δεδομένων που παρέχει τη δυνατότητα διόρθωσης και επισημείωσης «λαθών» μεταφορτώνονται τα κείμενα από τους/τις εκπαιδευτικούς, με σκοπό την ανατροφοδότηση και των δύο αλλά ταυτόχρονα και τη δημιουργία ενός ΜΣΚ που μπορεί να εμπλουτίζεται συνεχώς και να υπάρχει διαθέσιμο για μελλοντική χρήση (Mukherjee και Rohrbach 2006 · Gilquin και Granger 2010).

Ως απόρροια των προαναφερθέντων, τα ΜΣΚ μπορούν να αξιοποιηθούν από τους/τις μαθητές/τριες στην τάξη ή κατά τη μελέτη τους στο σπίτι. Μέσα από τη μελέτη και τη σύγκριση διαφορετικών αυθεντικών κειμένων, κάνουν εξάσκηση της παρατηρητικότητάς τους και καλλιεργούν τη γλωσσική τους επίγνωση. Ακόμη, μπορούν να εξασκούνται στους διαφορετικούς τρόπους παραγωγής νοήματος ενώ παράλληλα μπορούν να παρατηρούν τη συχνότητα των τρόπων που επιλέγουν, επιτυγχάνοντας με αυτόν τον τρόπο μεγαλύτερη αυτοπεποίθηση και εξελίσσοντας τους εαυτούς τους σε αυτόνομους ομιλητές/τριες της προς εκμάθηση γλώσσας. Με πιο απλά λόγια, αντί να υπάρχει καταιγισμός κανόνων που να πρέπει να απομνημονεύσουν και να αναπαράγουν οι μαθητές/τριες, τους δίνεται η δυνατότητα μέσα από τη συγκριτική μελέτη διαφορετικών κειμένων αυθεντικού λόγου να εντοπίζουν οι ίδιοι/ίδιες τον τρόπο που χρησιμοποιείται η γλώσσα και να εξυπηρετούν έτσι τις επικοινωνιακές τους ανάγκες (Johns 1991 · Chambers 2010 · Cheng 2010 · Sripicharn 2010). Τα ΣΜΚ μπορούν να χρησιμοποιηθούν ακόμη και στην επιμόρφωση εκπαιδευτικών που πρόκειται να διδάξουν μια γλώσσα ως Γ2 ή Γ1. Η βιβλιογραφία που ασχολείται με τέτοιου είδους μελέτες καταγράφει πως αυτό το υλικό είναι ένας από τους πιο σημαντικούς τρόπους που παρέχει εξοικείωση στους φοιτητές/τριες ή στους προς εκπαίδευση εκπαιδευτικούς, τονίζοντας τις πραγματικές ανάγκες και τις δυσκολίες που έρχονται αντιμέτωποι οι μαθητές/τριες (O’Keeffe και Farr 2003 · Römer 2008). Αν ληφθεί υπόψη πως ο εντοπισμός των αναγκών αλλά και η αξιολόγηση των γλωσσικών παραγόμενων αποτελούν βασικές πρακτικές για όσους διδάσκουν μια γλώσσα, κρίνεται απαραίτητο να είναι σε θέση να αναλύουν και να ερμηνεύουν τα «λάθος» σημεία των μαθητικών γραπτών (Αρχάκης 2011). Τα ΜΣΚ μπορούν να χρησιμοποιηθούν και για την παραγωγή λεξικών (Gillard και Gadsby 1998), για προγράμματα σπουδών, για γλωσσικά εγχειρίδια και για υλικό που σχετίζεται με τη γλωσσική διδασκαλία. Αυτό σημαίνει πως οι δυσκολίες και οι ιδιαίτερες ανάγκες των μαθητών/τριών που προκύπτουν στα ΜΣΚ, πρόκειται να επηρεάσουν τον επιτονισμό που δίνεται σε κάποια γλωσσικά φαινόμενα, τη σειρά διδασκαλίας, τα κείμενα που θα παρουσιαστούν για τη διδασκαλία και τέλος οι δραστηριότητες που θα κληθούν να κάνουν οι μαθητές/τριες.

Οι άνωθεν προτάσεις για την αξιοποίηση των ΜΣΚ προαπαιτούν ωστόσο διαθέσιμους πόρους, όπως χρηματοδότηση για την δημιουργία και τη συντήρηση των ΣΚ, το κατάλληλο λογισμικό και το σπουδαιότερο τεχνικές γνώσεις για την εξοικείωση χρήσης τους από τους/τις εκπαιδευτικούς. Επιπροσθέτως, η επιλογή, η προετοιμασία του υλικού καθώς και η επεξεργασία του στην σχολική αίθουσα μπορεί να αποτελέσουν χρονοβόρα διαδικασία

(Granger 2002 · O’Keeffe και Farr 2003 · Gilquin και Granger 2010). Παρόλο που όλα αυτά μπορούν να αποτελέσουν ανασταλτικό παράγοντα για τη χρήση ΜΣΚ για διδακτικούς σκοπούς, θεωρείται πως τα οφέλη από ένα τέτοιο εγχείρημα είναι πιο σημαντικά από τη δυσκολία τους.

5. ΓΛΩΣΣΙΚΟ ΛΑΘΟΣ

5.1. Ορισμός και προβληματισμοί

Ως γλωσσικό λάθος μπορεί να οριστεί κάθε απόκλιση από τη γλωσσική νόρμα, την οποία η κοινότητα θέτει για πρότυπο και την περιβάλλει με το κύρος του κανόνα. Εξαρχής με αυτόν τον ορισμό δημιουργούνται θέματα που κάνουν φανερό ότι παρά την αυστηρή αξιολογική θέση που έχουμε σε σχέση με τα γλωσσικά λάθη, τα πράγματα εμφανίζουν σχετικότητα. Ως λάθος μπορεί να χαρακτηριστεί οποιαδήποτε απόκλιση που εμφανίζεται σε μία γλωσσική ποικιλία ή διάλεκτο. Επιπλέον, η νόρμα συγκροτεί μια τυποποιημένη γλωσσική μορφή που αποκρυσταλλώνεται μέσα από μια διαδικασία κωδικοποίησης της, στη γραμματική και τα λεξικά. Πάραυτα, ακόμα και αυτού του είδους η τυποποίηση μπορεί να διαφέρει ανάλογα με το θέμα και τους στόχους που θέτονται σε ένα εγχειρίδιο, λόγω χάρη εάν στόχος είναι η περιγραφή, θα διαφέρει το αποτέλεσμα με το αν ο στόχος είναι η ρύθμιση (Παπαναστασίου 2001, 199).

Εξάλλου, η μορφή που παρουσιάζει η γλώσσα στα λεξικά και τις γραμματικές αντιπροσωπεύει τη συγκεκριμένη χρονική περίοδο που αυτά συντάχτηκαν, επομένως, δε μπορούν να καταγράψουν τη γλωσσική αλλαγή που υφίσταται κάθε στιγμή σε όλα τα επίπεδα της γλώσσας. Παρουσιάζεται έτσι μια απόσταση μεταξύ της «ζωντανής» γλώσσας που χρησιμοποιείται καθημερινά στο λόγο και της στατικής εικόνας που εμφανίζει η γλώσσα στα εγχειρίδια. Μια εικόνα που σίγουρα δε μπορεί να περιλάβει τη ζωντάνια, την ποικιλία και την ασυνέχεια του προφορικού λόγου. (Σετάτος 1991, 30-31). Επομένως, πολλές μορφές και εκφορές που ειπώθηκαν, αν τις κρίνουμε από την οπτική των εγχειριδίων, θα εκτιμηθούν ως λάθη.

Η σχετικότητα αυτή, που γίνεται φανερή μόλις διερωτηθούμε για τις παραμέτρους που ορίζουν το γλωσσικό λάθος αλλά και από αυτό που εμείς οι ίδιοι στην καθημερινότητα μας κρίνουμε ως λάθος, συχνά απολυτοποιείται. Αυτό συμβαίνει γιατί όταν η νόρμα συνδυάζεται

με το κύρος του κανόνα και την ισχύ του προτύπου, παρουσιάζεται ως η μόνη σωστή μορφή γλωσσικά. Γι' αυτόν το λόγο υπάρχει η αίσθηση ότι η γραμματική και τα λεξικά καταγράφουν τη «σωστή» γλωσσική χρήση. Αναντίρρητα λοιπόν, αυτό σε σχέση με τη γλώσσα, δημιουργεί απόλυτες στάσεις. Από τη μια προκαλείται κοινωνικός ρατσισμός όσων δεν ταυτίζονται με την επίσημη γλώσσα, και από την άλλη δημιουργούνται κινήσεις καθαρισμού αφού η παρέκκλιση από τη νόρμα παίρνει εκτεταμένη μορφή, κινήσεις που συνοδεύονται από απόψεις για κίνδυνο αλλοίωσης είτε ακόμη και αφανισμού της εκάστοτε γλώσσας (Παπαναστασίου 2001, 199).

Για έναν εκπαιδευτικό στο πλαίσιο του μαθήματος της γλωσσικής διδασκαλίας, που καλείται να διδάξει τη νόρμα, ως παράγοντα κοινωνικής συνοχής, είναι απαραίτητο το γλωσσικό λάθος να εντοπίζεται και να διορθώνεται. Το σημαντικό ωστόσο είναι να γίνει αντιληπτό πως πολλά γλωσσικά λάθη ξεκινούν από μηχανισμούς του ίδιου του γλωσσικού συστήματος, κάτι που μπορεί να κατευθύνει την προσέγγιση των λαθών σε έναν εναλλακτικό τρόπο διδασκαλίας, ή ακόμη να περιορίσει τις απόλυτες αξιολογικές θέσεις απέναντι στο λάθος (Παπαναστασίου 2001, 200).

5.2. Ελληνική γλώσσα και ιστορική ορθογραφία

Η νέα ελληνική γλώσσα ανάγεται στη γραφή με σύστημα ιστορικής ορθογραφίας. Στην αρχαία ελληνική και κυρίως στην αττική διάλεκτο, που από το 403 π.Χ καθιερώθηκε το ιωνικό αλφάβητο στην ορθογραφία, τα φωνήματα και τα γραφήματα βρίσκονταν σε άμεση σχέση. Στη νέα ελληνική όμως, δεν παρατηρείται αυτή η αμφιμονοσημαντικότητα, εξαιτίας των φωνητικών εξελίξεων που συνέβησαν. Η φωνητική ταύτιση των γραφημάτων <ι>, <η>, <υ>, <ει>, <οι>, <υι> σε [i] (ιωτακισμός), των <ε> και <αι> σε [e] και των <ο> και <ω> σε [o], σε συνδυασμό με την απλοποίηση των διπλών συμφώνων, αποτέλεσαν την αιτία για την ανατροπή της αρχής της αμφιμονοσημαντικότητας (Πετρούνιας 1984, 239-245 στο Παπαναστασίου 2001, 194).

Η πολυπλοκότητα των ορθογραφικών κανόνων εξαιτίας της αναντιστοιχίας στην προφορά και τη γραφή, δυσκολεύει πολύ τους/ τις ομιλητές/τριες και τους/τις μαθητές/τριες στον ορθό τρόπο γραφής της νέας ελληνικής γλώσσας. Η συνειδητοποίηση του προβλήματος φέρνει στην επιφάνεια συζητήσεις και τρόπους απλοποίησης, που χωρίς να υποσκάπτουν την παράδοση, θα διευκολύνουν την εκμάθηση των ορθογραφικών αναπαραστάσεων. Ωστόσο ο

συμβολικός χαρακτήρας της γραφής και της ορθογραφίας γενικότερα έχει άμεση συνάφεια με την εθνική ταυτότητα των Ελλήνων και αποτελεί για πολλούς, τρόπο σύνδεσης με το ένδοξο παρελθόν και συχνά ακόμη δυσκολότερη οποιαδήποτε προσπάθεια για αλλαγή (Αναγνωστόπουλος 1964, 107-111 στο Παπαναστασίου 2001, 195 · Κακριδή - Φερράρι 2008 · Τζωρτζάτου κ.α. 2019, 377).

Η μέχρι τώρα πορεία της γλώσσας μας έχει περάσει, ύστερα από διαπραγματεύσεις, σε δύο μεγάλες φάσεις αλλαγών. Οι αλλαγές αυτές είχαν κυρίως χαρακτήρα ιδεολογικοπολιτικό και πραγματοποιήθηκαν τον 20^ο αιώνα. Η πρώτη ταυτίζεται με την πτώση των συνταγματαρχών ενώ η επόμενη με την άνοδο της κεντροαριστεράς στην εξουσία. Πιο συγκεκριμένα, οι πρώτες προσπάθειες έγιναν από τον Μανόλη Τριανταφυλλίδη με απλουστεύσεις στην ορθογραφία και ενοποίηση των πολλαπλών γραφών που υπήρχαν στην εποχή του, ώσπου φτάσαμε στην εκπαιδευτική μεταρρύθμιση του 1976 για να καθιερωθεί η δημοτική ως επίσημη γλώσσα του ελληνικού κράτους. Η αλλαγές αφορούσαν την κατάργηση της διαφορετικής ορθογράφησης =, στην οριστική και την υποτακτική σε ενεργητική και παθητική φωνή (πχ. Να λύσει, να λυθεί), την ενοποίηση της ορθογράφησης παραθετικών επιθέτων και επιρρημάτων σε –ο (πχ. νέος-νεότερος, σοφός - σοφώτατα), και τέλος την απλούστευση δάνειων λέξεων (πχ. Τρένο- τραίνο).⁴ Το 1982 ήταν η δεύτερη μεγάλη φάση αλλαγών με την κατάργηση του πολυτονικού συστήματος και την καθιέρωση του μονοτονικού⁵ (Τριανταφυλλίδης [1941] 1988, 405-407 στο Παπαναστασίου 2001, 195 · Κακριδή-Φερράρι 2008, 369-373 · Παπαναστασίου 2008, 77-78 · Τζωρτζάτου κ.α. 2019, 379-380).

Οι άνωθεν μεταρρυθμίσεις αφορούν το μερικό επανεγγραφισμό της ελληνικής γλώσσας και υλοποιήθηκαν για να επιτευχθεί ο εξορθολογισμός της γραφής με την προφορά, σε ένα πιο συγχρονικό επίπεδο. Ολικό επανεγγραφισμό γραπτής γλώσσας μπορούμε να πούμε πως αποτελεί η αλλαγή του αλφαβήτου της τουρκικής γλώσσας, όπου το λατινικό αλφάβητο αντικατέστησε το αραβικό (Παπαναστασίου 2008, 75–76). Οι αλλαγές αυτές γίνονται ύστερα από έντονο σκεπτικισμό και σε καμία περίπτωση δεν θα έπρεπε να αντιμετωπίζονται ως τροποποιήσεις που θίγουν την ουσία της γλώσσας, γιατί αναφέρονται μόνο στη γραπτή της αναπαράσταση. Δυστυχώς είναι αρκετοί αυτοί που ισχυρίζονται πως το εκπαιδευτικό σύστημα της Ελλάδας υστερεί μετά την κατάργηση του πολυτονικού συστήματος γραφής.

⁴ Για τις αλλαγές αναλυτικά βλ. Επίμετρο της *Νεοελληνικής Γραμματικής της Δημοτικής* [1941] 2002, 443-447.

⁵ Για τους κανόνες μονοτονικού συστήματος αναλυτικά βλ. Επίμετρο της *Νεοελληνικής Γραμματικής της Δημοτικής* [1941] 2002, 451-455.

Ωστόσο οι αποφάσεις αυτές μπορούν να δικαιολογηθούν εάν ειπωθούν από κάποιον/α απλό ομιλητή/τρια. Εάν αυτό γίνεται από επιστήμονες που έχουν γνώση των πραγματικών διαστάσεων του φαινομένου, τότε είναι άκρως αδικαιολόγητη (Παπαναστασίου 2001). Η εμμονή στην ιστορικότητα της γλώσσας σύμφωνα με την Καραντζόλα (1999, 815-824) γίνεται εντονότερη σε γλώσσες όπως η ελληνική, που είναι πιο ασθενείς μεγαλοποιώντας το παρελθόν. Δε συμβαίνει όμως το ίδιο και με τις ισχυρές γλώσσες, που λόγω έλλειψης ιστορικότητας προβαίνουν πιο εύκολα σε ορθογραφικές μεταρρυθμίσεις με σκοπό να ενισχύσουν το κύρος και την ευκολότερη διάδοση της γλώσσας τους σε όλα τα έθνη. Οι όποιες λοιπόν αλλαγές για τον εξορθολογισμό της γραφής με τη μορφή επανεγγραφισμού, είναι μέρος του γλωσσικού προγραμματισμού της χώρας αφού η λήψη των αποφάσεων γίνεται από εκπροσώπους του κράτους για συγκεκριμένους σκοπούς.

5.3 Κατηγορίες γλωσσικών λαθών

Για να μπορέσουν οι ερευνητές/τριες να εντάξουν σε μια κατηγορία τα λάθη, υποχρεούνται με βάση το επίπεδο ανάλυσης της γλώσσας να δημιουργήσουν μια ποικιλία από κατηγορίες λαθών. Ο Παπαναστασίου (2001, 200-202), ακολουθεί την εξής κατηγοριοποίηση:

1. λάθη στο γλωσσικό σύστημα:

Λάθη στο *φωνητικό/ φωνολογικό επίπεδο* : π.χ. η προφορά της λέξης διάρρηξη ως [djariksɪ] αντί του [diariksɪ] που είναι το κανονικό. Η προφορά αποτελεί προσαρμογή στις άλλες λέξεις που προφέρονται ως [dja]: π.χ. διαβάζω [djavazo].

Λάθη στο *μορφολογικό επίπεδο*: π.χ. η διατήρηση της εσωτερικής αύξησης στον τύπο της προστακτικής διέγραψε, αντί του διάγραψε που είναι ορθό, με την επίδραση του γ' προσώπου του αορίστου διέγραψε · ή, στο επίθετο διεθνής, η δημιουργία τύπου γενικής του διεθνή αντί του διεθνούς που είναι ορθό, με την επίδραση της γενικής ουσιαστικών όπως ο μαθητής/ του μαθητού.

Λάθη στο *συντακτικό επίπεδο*: π.χ. Η εκφορά «μετά Χριστού» αντί του ορθού «μετά Χριστόν» ύστερα από τη επίδραση της εκφοράς προ Χριστού · ως λάθη μπορούν επιπλέον να εκληφθούν και οι εκφορές *πιο ανώτερος, πιο καλύτερος* – που ανήκουν στο μορφοσυντακτικό επίπεδο – που υπάρχουν πλεονασματικά δύο μορφήματα (πιο, -τερος) για τη δήλωση του συγκριτικού βαθμού.

Λάθη στο *σημασιολογικό επίπεδο*: π.χ. η χρήση της λέξης ευάριθμος με τη σημασία *πολύαριθμος*, «πολύς στον αριθμό» αντί της κανονικής σημασίας *ολιγάριθμος*, «εύκολος να αριθμηθεί».

Λάθη στο *λεξιλογικό επίπεδο*: π.χ. η χρήση της λέξης ανθηρόστομη στη θέση του αθυρόστομη· η χρήση του αποθανατίζω αντί απαθανατίζω.

2. Λάθη ως προς τη χρήση της γλώσσας:

Στην ενότητα αυτή εντάσσονται λάθη του προφορικού κυρίως λόγου (και πιο σπάνια του γραπτού π.χ. σε μια επιστολή), όπου για την κατανόηση τους προαπαιτείται η εμπλοκή κοινωνικών παραγόντων, και πιο συγκεκριμένα η αλληλεπίδραση πομπού και δέκτη. Ως λάθη αναγνωρίζονται είτε η άγνοια ανάμειξης στοιχείων ενός γένους/ είδους με ένα άλλο (π.χ. αφήγηση - επιχειρηματολογία), είτε η ανάμειξη διαφόρων επιπέδων ύφους (λαϊκός λόγος σε επιστημονικό κείμενο) είτε η παράβαση κανόνων γλωσσικής συμπεριφοράς σε συγκεκριμένα κοινωνικά συμφραζόμενα (π.χ. η χρήση ενικού όταν απευθυνόμαστε σε ανώτερο στέλεχος στην εργασία). Η εμπλοκή του κοινωνικού παράγοντα μπορεί να εντοπιστεί και σε λάθη υπερδιόρθωσης (όταν συμβαίνουν μη εσκεμμένα, με σκοπό να αποδώσουν προσωπικό ύφος): π.χ. εξ επί τούτου, λεπτά αντί για λεφτά (=χρήματα).

3. Λάθη που σχετίζονται με τη γραπτή μορφή της γλώσσας

Στην κατηγορία αυτή εντάσσονται κυρίως τα ορθογραφικά λάθη, που απασχολούν μεγάλο μέρος της γλωσσικής διδασκαλίας και συχνά υπερτονίζονται. Αυτό βέβαια, δε σημαίνει πως η ορθογραφία είναι ένα πεδίο που δε χρειάζεται εξάσκηση. Μεγάλο μέρος των ορθογραφικών λαθών γίνονται εξαιτίας της «ιστορικής ορθογραφίας» είτε ως ατελούς εκμάθησης είτε ως μη συνειδητοποίησης των κανόνων της (π.χ. λύνετε αντί λύνεται / αμφισβητείστε αντί αμφισβητήστε). Λάθη που εντάσσονται σε αυτή την κατηγορία είναι και τα στοιχεία προφορικότητας στο γραπτό λόγο.

Παρόμοια είναι και η ταξινόμηση που έκαναν οι Κουτσογιάννη και Παπαδοπούλου (2007, 176), με κατάταξη λαθών σε πέντε γενικές κατηγορίες και αντίστοιχα η κάθε μία από αυτές σε άλλες υποκατηγορίες. 1). Ορθογραφία (κλιτικό, παραγωγικό, λεκτικό μόρφημα), 2) τονισμός (λάθος τονισμός, άτονη λέξη), 3) μορφοσύνταξη (χρήση άρθρων, χρήση κλιτικών, σχέσεις συμφωνίας, μορφολογία ουσιαστικού και ρήματος, χρήση συμπληρωματικών

δεικτών, σειρά όρων),4) σημασιο-σύνταξη (χρήση προθέσεων, συνδέσμων, χρήση *θα* και *να*, χρήση κειμενικών δεικτών) και τέλος, 5) λάθη στο λεξιλόγιο σε επίπεδο σημασίας και μορφής.

Οι Ιορδανίδου και Αμπάτη (2007) όπως επίσης και οι Αναστασιάδη- Συμεωνίδη, Βλέτση, Μητσιάκη, Μποζονέλος & Χούμα (2008), χρησιμοποίησαν ένα σύστημα κατηγοριοποίησης λαθών βάσει των διαφόρων επιπέδων ανάλυσης της γλώσσας.

Από την άλλη, οι Αντωνοπούλου, Βελετόπουλος, Καρακύργιου, Μουμτζή και Παναγιωτίδου (2006, 21, 26-27), για την ανάλυση λαθών γραπτού λόγου στις εξετάσεις ελληνομάθειας του ΚΕΓ έκαναν ταξινόμηση των λαθών σε 14 κατηγορίες βάσει των μερών του λόγου της παραδοσιακής γραμματικής : 1) Άρθρο, 2) Ουσιαστικό, 3) Επίθετο 4) Αντωνυμίες, 5) Ρήμα, 6) Επίρρημα, 7) Πρόθεση 8) Σύνδεσμοι, 9) Ορθογραφία, 10) Στίξη, 11) Τονισμός, 12) Λεξιλόγιο, 13) Πλάγιος Λόγος και 14) Σύνταξη.

Ξένοι μελετητές όπως ο Dulay, Burt και Krashen (1982) κατηγοριοποίησαν τα λάθη α) σε γλωσσολογικές κατηγορίες (πχ. χρόνο, όψη, γένος) και β) στον τρόπο πραγμάτωσης τους (πχ. προσθήκη, παράλειψη, κτλ.). Η κατηγορίες πραγματοποιούνται σε ιεραρχία που περιλαμβάνει τρία στρώματα :τομέας λαθών (error domain), κατηγορία λαθών (error category) και γραμματική κατηγορία της λέξης (word category). Πρακτικά δηλαδή είναι σαν να πρότειναν δύο ξεχωριστές κατηγοριοποιήσεις (Τάντος κ.α. 2015, 225-126).

Ο Τοπο (2003), πραγματοποιεί μια ταξινόμηση η οποία καταφέρνει να συνδυάσει τις δύο προηγούμενες, μιας και απαρτίζεται από τη μία με γλωσσολογικές κατηγορίες με την ακόλουθη ιεράρχηση: 1) επίπεδο (level): φωνολογία, γραφολογία, γραμματική, λεξιλόγιο, κείμενο, λόγος, 2) τάξη (class): ουσιαστικό, ρήμα, επίθετο, επίρρημα, πρόθεση, σύνδεσμος, προσδιοριστής, 3) διάκριση (rank): μόρφημα, λέξημα, φράση, πρόταση, περίοδος, 4) σύστημα (system): χρόνος, αριθμός, φωνή, μετρησιμότητα, σαφήνεια (sense-relation), τοποθέτηση (collocation) και από την άλλη βάσει των ειδών τροποποίησης της επιφανειακής δομής: προσθήκη, παράληψη, παραμόρφωση, ανάταξη, σύγχυση.

Η Granger (2003, 467-469 στο Τζιμώκας 2018, 44) στη συνέχεια στηριζόμενη στην άποψη της Τοπο και του James (1998, 106-113), εισάγει και αυτή έναν νέο κώδικα ταξινόμησης λαθών που συνδυάζει κριτήρια κατηγοριοποίησης, που διαρθρώνονται ιεραρχικά σε τρία επίπεδα: α) επίπεδο λάθους β) κατηγορία λάθους και γ) κατηγορία λέξης ή κατηγοριοποίηση λέξης ως προς τα μέρη του λόγου. Πρόκειται για ταξινόμηση με περιγραφικό κυρίως

χαρακτήρα, που δεν περιλαμβάνονται διακρίσεις ερμηνευτικού τύπου, όπως λάθος – γλωσσική παραδρομή και διαγλωσσικά – ενδογλωσσικά λάθη, γιατί αυτές είναι δύσκολο να αποδοθούν σε κάθε περίπτωση με ακρίβεια επομένως είναι καλύτερο να διερευνώνται σε μεταγενέστερο στάδιο της ανάλυσης των λαθών.

5.3.1. Ορθογραφικά λάθη

Τα ορθογραφικά λάθη ανήκουν σε μία από τις κατηγορίες των γλωσσικών λαθών. Σύμφωνα με τον ορισμό του Παπαναστασίου (2008, 87), το ορθογραφικό λάθος ορίζεται ως «η αστοχία του συντάκτη του κειμένου στην επιλογή του κατάλληλου γραφήματος για την απόδοση ενός φωνήματος και η χρήση στη θέση του κάποιου άλλου, που έχει την ίδια αξία».

Παρόλο που η δυσκολία των μαθητών/τριών στην εκμάθηση ορθογραφικών κανόνων είναι γνωστή, η αστοχία τους στη σωστή ορθογραφία της λέξης αντιμετωπίζεται ως γλωσσικό έλλειμμα. Το μήνυμα που μεταφέρεται από ένα μη ορθογραφημένο κείμενο είναι η άγνοια και ως επί το πλείστον η παραβίαση των γλωσσικών και ορθογραφικών κανόνων.

Σύμφωνα ωστόσο με έρευνες των Clark και Ivanic (1997, 199-200), δεν είναι απαραίτητο ο μαθητής/τρια που κάνει ορθογραφικά λάθη, να αγνοεί πάντα τους κανόνες. Πολλά άτομα εμφανίζουν δυσκολία στην απομνημόνευση λεξικών εικόνων, κάτι που προκαλεί δυσκολία κατά την ορθογραφία. Στην έρευνα τους, επισημαίνουν πέραν των εγκεφαλικών λειτουργιών και της οπτικής μνήμης που δείχνουν πως επηρεάζουν την ορθογραφική ικανότητα των ομιλητών/τριών, και τον ρόλο που έχει η ακουστική και η κιναισθητική μνήμη στην όλη διαδικασία. Με λίγα λόγια, η ικανότητα ενός μαθητή/τριας να παράγει σωστή ορθογραφία, είτε πρόκειται για μονόγλωσσο είτε για δίγλωσσο άτομο, καθορίζεται από παράγοντες που δεν έχουν πάντα άμεση σχέση με τη γενικότερη γνώση του μαθητή/τριας για τη γλώσσα.

Στις περιπτώσεις βέβαια που δεν πληρούνται οι ορθογραφικοί κανόνες λόγω κάποιας δυσκολίας, αυτό μπορεί να προσληφθεί μόνο ως έλλειμμα. Ο Παπαναστασίου (2008, 87), παρόλο που γνωρίζει ότι η εκμάθηση των ορθογραφικών κανόνων είναι δύσκολη, υποστηρίζει πως η ασυμβατότητα με την ορθογραφική νόρμα σε ένα γραπτό κείμενο, δεν ταιριάζει με τον κοινωνικό χαρακτήρα της ορθογραφίας. Η ορθογραφία, ως αναπαράσταση της νόρμας, αποτελεί μέρος της εκπαιδευτικής διαδικασίας και επομένως όσοι/ες επιζητούν πρόσβαση σε κοινωνικά επιτεύγματα, πρέπει να ακολουθήσουν τους κανόνες της.

Αυτός είναι και ο λόγος που πρέπει οι εκπαιδευτικοί να διορθώνουν τα ορθογραφικά λάθη των μαθητών/τριών. Σύμφωνα με τον Χαραλαμπίδη (2001, 197), οι απαραίτητοι ορθογραφικοί κανόνες εκμάθησης αφορούν το γένος και την κλίση των ονομάτων. Μια δυσκολία που αντιμετωπίζουν κυρίως οι αλλόγλωσσοι μαθητές/τριες. Εξάλλου δεν είναι λίγες οι έρευνες που έγιναν σε μαθητές/τριες που μαθαίνουν την ελληνική ως δεύτερη γλώσσα που επιβεβαιώνουν πως τα περισσότερα λάθη προκύπτουν από την ορθογράφιση των λέξεων. Το ότι είναι σχετική η φύση των ορθογραφικών κανόνων κάθε γλώσσας είναι επιβεβαιωμένο, ωστόσο η γνώση τους είναι μείζονος σημασίας τόσο σε εκπαιδευτικό όσο και σε κοινωνικό επίπεδο.

Συμπερασματικά, τα ορθογραφικά λάθη είναι αποκλίσεις από τον κανόνα. Παρά όμως το συμβατικό χαρακτήρα που έχει η ορθογραφία, τα λάθη αυτά αξιολογούνται όπως και τα άλλα γλωσσικά λάθη (φωνολογικά, μορφολογικά κα.).

5.3.2 Κατηγοριοποίηση Ορθογραφικών λαθών

Η ταξινόμηση των ορθογραφικών λαθών έχει μεγάλη πορεία στην ελληνική γλώσσα ωστόσο οι δημοσιευμένες προσεγγίσεις είναι λίγες, ενώ ακόμη λιγότερες είναι οι προσεγγίσεις με συστηματικότητα. Μια από τις σχετικά πρόσφατες και αξιόλογες κατηγοριοποιήσεις που υλοποιήθηκαν τα τελευταία χρόνια είναι η μελέτη των Κουτσογιάννη και Παπαδοπούλου (2007) οι οποίοι ομαδοποίησαν τα λάθη σε α) ορθογραφικά, β) μορφοσυντακτικά, γ) τονικά, δ) σημασιοσυντακτικά, ε) λεξιλογικά. Κάθε κατηγορία αποτελείται από άλλες υποκατηγορίες. Συγκεκριμένα θα αναφερθούμε στην υποκατηγορία των ορθογραφικών που είναι και αυτή που μας ενδιαφέρει. Τα ορθογραφικά λάθη λοιπόν αφορούν α) τα κλιτικά επιθήματα, β) τα παραγωγικά επιθήματα γ) και τα λεξικά μορφήματα. Τα κλιτικά μορφήματα αφορούν καταλήξεις από τα κλιτά μέρη του λόγου που χρησιμοποιούνται για να δηλώσουν γραμματική σχέση, όπως γένος, αριθμό, πτώση, πρόσωπο, χρόνο ενώ τα παραγωγικά μορφήματα αφορούν τα προθήματα και τα επιθήματα που χρησιμοποιούνται όταν παράγεται νέα λέξη. Τα λεξικά μορφήματα αφορούν τα θέματα των κλιτών λέξεων, όπως το θέμα από ρήματα, ουσιαστικά, επίθετα καθώς και άκλιτες λέξεις όπως επιρρήματα (Κουτσογιάννης και Παπαδοπούλου 2007,176). Ωστόσο δεν προσφέρουν επιπλέον ανάλυση ή εμβάθυνση σε πιθανές νοηματικές σχέσεις για τις αστοχίες των ορθογραφικών λαθών (Πρωτόπαπας και Σκαλούμπακας 2010, 181-182).

Μια πιο συστηματική και θεωρητικά βασισμένη ταξινόμηση ορθογραφικών λαθών είναι αυτή των Πρωτόπαπα και Σκαλούμπακα (2010,182-196) που δημιούργησαν έξι (6) γενικές κατηγορίες και κάθε μία από αυτές διακρίνεται σε περαιτέρω υποκατηγορίες λαθών.

Φωνολογικά: Λάθη στη γραφή των λέξεων που δε διατηρούν τη σωστή απόδοση των φθόγγων και προκαλούν αλλοίωση στη φωνολογική ταυτότητα των λέξεων.

Υποκατηγορίες:

- αντικατάσταση γραφήματος (*συνεσφίματα/συναισθήματα BGs001⁶*),
- παράλειψη γραφήματος (*χήματα/χρήματα*),
- προσθήκη γραφήματος (*αντινί/αντί*), αντιμετάθεση διπλανών γραφημάτων (*γεωργοί/γεωργοί*),
- αντιστροφή δίψηφου συμφώνου (*πμαίνει/μπαίνει*),
- απλοποίηση δίψηφου φωνήεντος (*ατί/αντί*), παράλειψη διαλυτικών (*προιόντα/προϊόντα*),
- παράλειψη συλλαβής (*κουνίζω/κουδουνίζω*).

Γραμματικά λάθη: Έχουν άμεση εξάρτηση από το μορφολογικό τύπο της λέξης. Προϋπόθεση για την περαιτέρω ανάλυση τους είναι αν πρόκειται για κλιτό ή άκλιτο τύπο.

Υποκατηγορίες:

- Γραμματικό λάθος σε κλιτή λέξη (*αμάξη, ABm003*),
- Γραμματικό λάθος σε άκλιτη λέξη (*ακριβός, AGm008*)
- Γραμματικό λάθος σε κλιτικό πρόθημα (*αίδωσα* αντί *έδωσα*)

Ετυμολογικά - οπτικά λάθη (ιστορικής ορθογραφίας): Στη φωνογραφημική διαφάνεια⁷ η διαφορά των συμφώνων με τα φωνήεντα είναι μεγάλη, και επομένως η ορθογραφία τους γίνεται δυσκολότερη. Έτσι οι υποκατηγορίες αφορούν την εμφάνιση συμφώνου ή φωνήεντος στο θέμα της λέξης.

Υποκατηγορίες:

⁶ Όσες κατηγορίες καλύπτονταν από τα δικά μας γραπτά χρησιμοποιήθηκαν ως παραδείγματα μαζί με τα αντίστοιχα γραπτά που ανιχνεύτηκαν (για την κωδικοποίηση των γραπτών βλέπε *παρ. 6.8.*), τα υπόλοιπα παραδείγματα δανείστηκαν από τους Πρωτόπαπα και Σκαλούμπακα (2010, 196-197)

⁷ Με τον όρο διαφάνεια σε ένα ορθογραφικό σύστημα εννοείται η συνέπεια και η συστηματικότητα που απεικονίζει τους φθόγγους της γλώσσας σε γράμματα ή συνδυασμούς γραμμάτων (γραφήματα), (Μουζάκη και Πρωτόπαπας 2010, 541).

- Ετυμολογικό - οπτικό λάθος σε φωνήεν. Ένας φωνηεντικός φθόγγος αποδίδεται με λάθος γράφημα του ίδιου φθόγγου (*αρνιτικά*, BBm008),
- ετυμολογικό οπτικό λάθος σε σύμφωνο, όπου ένας συμφωνικός φθόγγος αποδίδεται με λάθος γράφημα του ίδιου φθόγγου (*μαζεφτίκαμε*, GBm003).
- Θεματικό λάθος σε περίπτωση κανόνα. Πρόκειται για λάθη στο θέμα των λέξεων που ελέγχεται από αυστηρά καθορισμένους κανόνες παραγωγικής μορφολογίας που διδάσκονται συστηματικά. Αυτά τα λάθη ίσως διαφέρουν από άλλα λάθη ιστορικής φύσεως γιατί η αποφυγή τους δεν εξαρτάται μόνο από την απομνημόνευση μιας συγκεκριμένης λέξης και των παραγώγων της, αλλά υποβοηθείται από άλλες λέξεις με διαφορετικά θέματα που όμως απαρτίζουν παραγωγική οικογένεια βάση του παραγωγικού επιθήματος και όχι βάση του θέματος. Τα λάθη αυτά δεν αντιμετωπίζονται ως γραμματικά επειδή τα παραγωγικά επιθήματα αλλάζουν το θέμα και έχουν νοηματικές συνέπειες για τη λέξη, ενώ ακολουθεί γραμματική (κλιτική) κατάληξη στην οποία το λάθος θα ήταν γραμματικού τύπου. Στην κατηγορία αυτή ανήκουν τα ρήματα σε -ίζω, -ώνω, -αίνω, -εύω, καθώς και λάθη ουσιαστικών σε -ικός, -ίδα, -ένιο, -ίτης (*αντιμετώπισης*, GBm011, *συνυθίσαμε*, AGm002, *γεμίζω*, *απλώνω*, *γυμναστική*, *κουρέβω*).
- θεματικό λάθος σε εξαίρεση κανόνα. Αφορά σε λάθη στο θέμα λέξεων των οποίων η ορθογραφία ρυθμίζεται από τις εξαιρέσεις των κανόνων της προηγούμενης περίπτωσης. Για παράδειγμα, αφορά σε λάθη στο θέμα «ε» των ρημάτων *πλένω*, *δένω*, *μένω* που αποτελούν τις εξαιρέσεις του κανόνα ο οποίος ορίζει πως τα ρήματα που λήγουν σε -αίνω γράφονται με αι (*πλαίνομαι/πλένομαι*), λάθη στο «ει» του ρήματος *δανείζω*, που αποτελεί την εξαίρεση του κανόνα ο οποίος ορίζει πως τα ρήματα που λήγουν σε -ίζω γράφονται με ι (*δανίζουν*)
- παραβίαση ουρανικοποίησης. Λάθη λόγω άγνοιας ή ελλιπούς κατανόησης των συμβάσεων της ορθογραφίας για τη γραφή των ουρανικών συμφώνων και την εφαρμογή τους σε συγκεκριμένη λέξη με την παρουσία ή απουσία γραφήματος τύπου ι (*παχιές/παχές*, *κεργιά/κεριά*)

Τονικά λάθη: αναλόγως της παρουσίας ή της απουσίας τόνου προκύπτουν οι εξής υποκατηγορίες:

- Παράλειψη βασικού τόνου (*απο/από*),

- λάθος θέση τόνου (*συνέπης/συνεπής*),
- τόνος σε μονοσύλλαβη λέξη (*θά/θα*),
- παράλειψη δεύτερου τόνου (*βοήθεια μας/βοήθειά μας*),
- επιπρόσθετος τόνος (*επαγγέλματίας/επαγγελματίας*).

Λάθη στη στίξη: διαχωρίζονται σε βασικά (τελεία, κόμμα, θαυμαστικό, ερωτηματικό) και δευτερεύοντα (όλα τα υπόλοιπα) σημεία στίξης καθώς και σε λάθη απουσίας ή παρουσίας.

Υποκατηγορίες:

- Βασική παράλειψη,
 - περιττή προσθήκη,
 - αντικατάσταση βασικού σημείου,
 - αντικατάσταση δευτερεύοντος σημείου,
 - προσθήκη δευτερεύοντος σημείου στίξης,
 - παράλειψη δευτερεύοντος σημείου.
- Άλλα λάθη: Λάθος χωρισμένες λέξεις, λάθος ενωμένες, λάθος χρήση πεζού/κεφαλαίου, κενά μεταξύ συλλαβών, κενά εντός λέξης, παράλειψη λέξης ή φράσης, προσθήκη λέξης, αντικατάσταση λέξης με άλλη, λέξεις με κενά, περιττή χρήση διαλυτικών, καθρεπτική γραφή (3 αντί ε).

5.4. Μελέτες ανάλυσης λαθών στην Ελληνική ως Γ1 και Γ2

Η Ιγνατιάδου (2008, 17-20) σε μελέτη που έκανε σε εκθέσεις μαθητών/τριών από τα τετράδια εξετάσεων του Α΄ και Β΄ επιπέδου για την πιστοποίηση επάρκειας της ελληνομάθειας του ΚΕΓ, οδηγήθηκε στο συμπέρασμα πως το 17,69% των ορθογραφικών λαθών του Α΄ επιπέδου και το 11,98% του Β΄ επιπέδου είχαν σχέση με λάθη στα διαφορετικά επίπεδα φωνολογικής επίγνωσης, τη φωνημική και συλλαβική επίγνωση και την επίγνωση συλλαβικής δομής. Κυρίως ήταν λάθη στην αναντιστοιχία γραφήματος με φώνημα (πχ. νεβριασμένοι).⁸ Λάθη που γίνονται σε επίπεδο φωνολογικής επίγνωσης αποδεικνύουν τη δυσκολία που έχουν

⁸ Η έρευνα της Ιγνατιάδου (2008) εντάσσει αυτή τη κατηγορία λαθών στα «φωνολογικά λάθη», εννοώντας λάθη που διατηρούν τη φωνολογική ταυτότητα της λέξης όπως τα μελέτησε και ο Ροιπράδας (2001). Οι Πρωτόπαπας και Σκαλούμπακας (2010) αυτά τα λάθη τα κατατάσσουν στα ετυμολογικά και συγκεκριμένα στην κατηγορία «οπτικό λάθος σε σύμφωνο».

μαθητές/τριες με δεύτερη γλώσσα την ελληνική κατά τα πρώτα στάδια εκμάθησης στην ανάλυση των λέξεων σε φωνήματα γιατί προφέρονται μαζί με άλλα. Ωστόσο, η φωνημική ικανότητα (αντιστοιχία γραφής - γραφήματος), κατακτάται μόλις μάθουν να διαβάζουν. Εντόπισε επίσης λίγα λάθη στη συλλαβική επίγνωση, στο χωρισμό δηλαδή των συλλαβών (πχ. πεζμου), και λάθη επίγνωσης στη συλλαβική δομή, στη συνειδητοποίηση πως η συλλαβή αποτελείται από δύο δομικά τμήματα (πχ. για υτή). Τα χαμηλά ποσοστά λαθών σε αυτούς τους δύο τομείς εξηγούνται από το γεγονός ότι οι συλλαβές γίνονται πιο αντιληπτές από τους/τις μαθητές/τριες με αποτέλεσμα να μπορούν πιο εύκολα να προβούν σε ανάλυση των λέξεων σε συλλαβές και σε συλλαβικά τεμάχια συγκριτικά με τα φωνήματα. Γενικότερα, κατά τον Πορπόδα (2002, 270 στο Ιγνατιάδου 2008, 19), η φωνολογική επίγνωση παίζει καθοριστικό ρόλο στη μάθηση της ανάγνωσης και της ορθογραφημένης γραφής της ελληνικής γλώσσας, έπειτα από εξάσκηση.

Η Κατσιμαλή (2007, 142, 152-155) σε έρευνα που πραγματοποιήθηκε σε Έλληνες ομογενείς που είχαν κυρίαρχη γλώσσα τα αγγλικά παρουσίασαν σύγχυση στα όρια των λέξεων. Στην ίδια διαπίστωση καταλήγουν και οι Ιορδανίδου και Αμπάτη (2007), υποστηρίζοντας πως ορισμένοι τύποι απομνημονεύονται από τους/τις μαθητές/τριες με Γ2 την ελληνική ως προφορικές φόρμουλες, ενωμένες (πχ. μαρεσει αντί μ' αρέσει).

Από την έρευνα της Κατσιμαλή (2007, 132-134) σε ενήλικες μαθητευόμενους/ες φυσικούς ομιλητές/τριες της ελληνικής γλώσσας που ήταν απόφοιτοι/τες μέσης εκπαίδευσης εντοπίστηκαν λάθη λόγω ελλιπούς γνώσης της ιστορικής ορθογραφίας (λάθη οπτικής μνήμης). Το ποσοστό των 52,2% έφτασαν οι λάθος απαντήσεις στην προσπάθεια σύνδεσης παράγωγης λέξης της νέας ελληνικής με την αρχαία ρίζα του ρήματος προέλευσης (πχ. επανάληψη, κλίση). Επομένως, οι μαθητές/τριες μη φυσικοί ομιλητές/τριες της ελληνικής που δεν έχουν καμία γνώση πάνω στην αρχαία ελληνική γλώσσα, για να πετύχουν ορθογραφημένη γραφή πρέπει να απομνημονεύσουν αυτές τις λέξεις ως μεμονωμένα λήμματα.

Στην Ιγνατιάδου (2008, 20-24, 43-46), από την ίδια έρευνα της ανάλυσης ορθογραφικών λαθών στο επίπεδο Α' και Β' των εξετάσεων για πιστοποίηση ελληνομάθειας του ΚΕΓ προέκυψαν τα ποσοστά 26,19% και 32,03% που αφορούσαν λάθη στα κλιτικά και παραγωγικά (προθήματα και επιθήματα) μορφήματα. Το υψηλότερο ποσοστό αφορούσε λάθη στα κλιτικά μορφήματα (κατάληξη λέξης) πιθανόν εξαιτίας των πολλών γραμματικών τύπων. Τα περισσότερα μάλιστα λάθη παρουσιάστηκαν στα ρήματα λόγω της πολυμορφίας των

καταλήξεων, ακολούθησαν τα ουσιαστικά και λιγότερο τα επίθετα και οι αντωνυμίες αφού η χρήση τους ήταν περιορισμένη.

Στο ίδιο συμπέρασμα καταλήγει και η έρευνα της Ηλιοπούλου (2009, 270-276). Οι μη φυσικοί ομιλητές/τριες κρατούν υψηλά τα ποσοστά (97,71%) των ορθογραφικών λαθών σε λάθη που αφορούν την οπτική μνήμη. Το ίδιο και οι φυσικοί/ες ομιλητές/τριες. Λιγότερα λάθη εμφανίζονται στη γραφηματική απόδοση των φθόγγων των αλλοδαπών μαθητών/τριών εξαιτίας της φωνητικής – φωνολογικής παρεμβολής της μητρικής τους στην ελληνικής γλώσσας. Ελάχιστα ήταν και τα λάθη που σχετίζονταν με τη μεταφορά λάθος γραφήματος από τη Γ1 στην ελληνική. Επιπλέον, οι αλλοδαποί/ες μαθητές/τριες στην Α΄ και Β΄ Γυμνασίου συνέχισαν τον προφορικό λόγο με τον γραπτό κάτι που δεν συνέβη στη Γ΄ Γυμνασίου. Αναλυτικότερα, προβληματική περιοχή ήταν τα διπλά όμοια σύμφωνα (πχ μμ, λλ), τα διπλά σύμφωνα (κς αντί ξ, πς αντί ψ), τα ομόηχα σύμφωνα (γγ-γκ), και οι διαφορετικές μορφές του /i/, /e/, /o/.

Σε έρευνα των Φάκου κα. (2010, 297-302) αναλύθηκαν τα ορθογραφικά λάθη μαθητών/τριών της Γ΄, Δ΄ και ΣΤ΄ δημοτικού σχολείου μεταξύ μαθητών/τριών με δυσλεξία και τυπικούς μαθητές/τριες. Τα αποτελέσματα στους/στις τυπικούς/ες μαθητές/τριες έδειξαν πως τα περισσότερα λάθη ήταν γραμματικά, ακολούθησαν τα τονικά και κατ' ακολουθία τα ετυμολογικά – οπτικά λάθη ενώ τα φωνολογικά λάθη ήταν ελάχιστα.

Στα ίδια ακριβώς συμπεράσματα οδηγήθηκε και η έρευνα των Protopapas et al. (2013) που έγινε σε παιδιά Γ΄ και Δ΄ δημοτικού και Α΄ γυμνασίου. Τα γραμματικά λάθη σε τυπικούς μαθητές/τριες ήρθαν πρώτα στα αποτελέσματα, ακολούθησαν τα τονικά και έπειτα τα ετυμολογικά – οπτικά.

Ωστόσο, δύο χρόνια νωρίτερα η Μαυροειδή (2011) μελετώντας 20 γραπτά ενηλίκων και όχι παιδιών συμπέρανε πως τα περισσότερα λάθη εντοπίστηκαν στην κατηγορία οπτικών λαθών και ακολούθησαν τα γραμματικά σε πολύ μικρότερο ποσοστό.

Γενικότερα ο Αϊδίνης (2010, 164) σε έρευνα που πραγματοποίησε σε μαθητές/τριες δημοτικού (Α΄-ΣΤ΄) και ερεύνησε τις στρατηγικές κατάκτησης της γλώσσας που χρησιμοποιούν τα παιδιά σε φωνολογικό, μορφολογικό και ετυμολογικό επίπεδο οδηγήθηκε στο συμπέρασμα πως στις πρώτες τάξεις, οι μαθητές/τριες υποβοηθούνται από φωνολογικές και οπτικές στρατηγικές και όσο μεγαλώνουν και εξοικειώνονται με τις λέξεις χρησιμοποιούν περισσότερο μορφολογικές στρατηγικές πρώτα για την κατάκτηση του κλιτικού μορφήματος και στη συνέχεια του θεματικού μορφήματος. Παρόλα αυτά τα παιδιά

συναντούν μεγάλη δυσκολία στη γραφή στο να αντιληφθούν τη σταθερή θεματική γραφή των λέξεων και ενώ καταφέρνουν και βελτιώνουν τις μορφολογικές τους γνώσεις, δε συμβαίνει το ίδιο στην ορθογράφιση των λέξεων. Εκεί που καταλήγει εν τέλει είν2αι πως η ανάπτυξη των στρατηγικών κατάκτησης της ελληνικής γλώσσας δεν ολοκληρώνεται στο τέλος του δημοτικού αλλά συνεχίζει και σε μεγαλύτερες τάξεις.

6. ΕΠΕΞΕΡΓΑΣΙΑ ΜΑΘΗΤΙΚΟΥ ΣΩΜΑΤΟΣ ΚΕΙΜΕΝΟΥ ΜΑΘΗΤΩΝ ΓΥΜΝΑΣΙΟΥ ΓΙΑ ΤΗΝ ΑΝΑΛΥΣΗ ΟΡΘΟΓΡΑΦΙΚΩΝ ΛΑΘΩΝ

6.1. Ταυτότητα της έρευνας

Μετά την παρουσίαση των ελληνικών σωμάτων κειμένου και ειδικότερα των σωμάτων κειμένου μαθητών (Learner Corpora) στην ελληνική και στην αγγλική γλώσσα ως Γ1 και Γ2 και έπειτα από μια σύντομη αναφορά στην ελληνική ορθογραφία, τα γλωσσικά λάθη και προηγούμενες έρευνες που αφορούσαν τα ορθογραφικά λάθη των μαθητών/τριών, προχωράμε στην πραγματοποίηση της έρευνας.

Σε πρώτη φάση θα προβούμε στην κατασκευή ενός μαθητικού σώματος κειμένων από γραπτά μαθητών και μαθητριών της Α΄, Β΄ και Γ΄ γυμνασίου. Σκοπός της έρευνας είναι μέσα από το μαθητικό σώμα κειμένων που δημιουργήθηκε να γίνει μια πρώτη επεξεργασία για την εξόρυξη δεδομένων στα ορθογραφικά λάθη που παρατηρούνται κατά την παραγωγή γραπτού λόγου. Στη μελέτη αυτή πρέπει ωστόσο να τονίσουμε πως δε θα εστιάσουμε διεξοδικά στην ερμηνεία των ορθογραφικών λαθών, αλλά στην εξαγωγή γενικότερων συμπερασμάτων σε σχέση με το βαθμό κατάκτησης των ορθογραφικών κανόνων της ελληνικής γλώσσας σχετικά με τα κριτήρια που έχουμε ορίσει.

Στη συνέχεια, θα γίνει παρουσίαση της μεθοδολογίας που χρησιμοποιήθηκε για να πραγματοποιηθεί η έρευνα με σκοπό να επιβεβαιωθούν ή να διαψευστούν οι αρχικές υποθέσεις τις οποίες μελετήσαμε ως προς τρεις ανεξάρτητες μεταβλητές, φύλο/ ηλικία-τάξη/ καταγωγή. Παρακάτω θα αναφερθούν πιο αναλυτικά ο σκοπός και οι στόχοι της έρευνας καθώς και οι προβληματισμοί που προκύπτουν στο κομμάτι αυτό. Τέλος, θα υπάρξει ανάλυση των δεδομένων και θα διεξαχθούν τα τελικά συμπεράσματα.

6.2. Σκοπός και στόχοι

Έρευνες του Γκότοβου (1992, 75) που έγιναν σε μαθητές/τριες δημοτικού, έδειξαν πως η κατάκτηση και η εφαρμογή των ορθογραφικών κανόνων είναι σε άμεση συνάρτηση με την ηλικία των μαθητών/τριών, κυρίως στο Δημοτικό. Η γνώση δηλαδή των ορθογραφικών συμβάσεων φαίνεται πως με την πάροδο του χρόνου βελτιώνεται. Στο ίδιο συμπέρασμα κατέληξε και η Παπαδοπούλου - Μανταδάκη (2004, 47) αναφέροντας χαρακτηριστικά «είναι εμφανές ότι όσο μεγαλώνουν τα παιδιά και αποκτούν περισσότερη εμπειρία στην ανάγνωση και στην οπτική αναπαράσταση των λέξεων, κάνουν λιγότερα ορθογραφικά λάθη». Επομένως, η ηλικία αποτελεί σημαντική παράμετρο για την αντιμετώπιση των δυσκολιών της ορθογραφίας αφού η γνώση και η συνειδητή εφαρμογή των κανόνων ορθογραφίας είναι κάτι που κατακτάται σταδιακά.

Ωστόσο σύμφωνα με τη διεθνή βιβλιογραφία γενικά οι ορθογραφικές γνώσεις των μαθητών και μαθητριών είναι προσωρινές και άστατες, γιατί όταν τα παιδιά αλλάζουν βαθμίδα από το Δημοτικό στο Γυμνάσιο συνήθως τις «ξεχνούν» ή τις εγκαταλείπουν. Ενώ δηλαδή είχαν μάθει στο Δημοτικό να γράφουν ορθογραφήμενα ή σχεδόν ορθογραφήμενα, στο Γυμνάσιο παρατηρείται να κάνουν λάθη από απροσεξία ή από λήθη των ορθογραφικών κανόνων. Αυτό έχει ως αποτέλεσμα οι απόφοιτοι των ελληνικών σχολείων να νιώθουν ανασφάλεια, να μην έχουν σταθερότητα στους ορθογραφικούς κανόνες και να μη μπορούν να αποφύγουν τα ορθογραφικά λάθη όταν γράφουν ένα κείμενο. (Καράμπελα 2007, 151-152).

Όσον αφορά τους δίγλωσσους μαθητές/τριες ο Cummins (2005, 105) παρατήρησε πως για την κατάκτηση διακριτών δεξιοτήτων που σχετίζονται με την ορθογραφική κατάκτηση, η ηλικία αποτελεί σημαντική παράμετρο. Αυτό συμβαίνει γιατί οι δίγλωσσοι μαθητές/τριες προχωρούν τις τάξεις έρχονται σε επαφή με τη γλώσσα και τα κείμενα περισσότερο, επομένως κατακτούν και σε μεγαλύτερο βαθμό τις ορθογραφικές συμβάσεις της γλώσσας στόχου.

Άλλες προγενέστερες μελέτες (Trudgill, 1974 · Cheshire, 1978 στο Τζωρτζάτου 2012, 43), έδειξαν πως τα δύο φύλα παρουσιάζουν διαφορές ως προς τη χρήση της γλώσσας. Τα αποτελέσματα του Trudgill σε μελέτες που πραγματοποιήθηκαν στο Norits της Αγγλίας, έδειξαν πως οι άντρες ομιλητές με ευκολία μπορούν να παραβιάσουν τους τύπους της νόρμας. Αυτό συμβαίνει γιατί τείνει να προβάλλεται το καλυμμένο γόητρο ενός δυνατού κοινωνικά άντρα μέσα από συνδηλώσεις σκληρότητας και τραχύτητας. Αντιθέτως, οι γυναίκες τείνουν να ακολουθούν τη γλωσσική νόρμα με σκοπό να ενισχυθεί το κοινωνικό

τους γόητρο, εξαιτίας της κοινωνικής ανασφάλειας που νιώθουν. Σε παρόμοια συμπεράσματα κατέληξε και η έρευνα του Cheshire στο Reading της Αγγλίας, όπου διαπίστωσε πως σε επίσημες περιστάσεις τα κορίτσια συμπορεύονται περισσότερο με τη νόρμα απ' ότι τα αγόρια.

Πάνω σε αυτά τα δεδομένα προηγούμενων ερευνών που αφορούν την ηλικία/τάξη, το φύλο και την καταγωγή, δόθηκε η εφόρμηση ώστε να δημιουργηθεί ένα μαθητικό σώμα κειμένου με γραπτά μαθητών και μαθητριών Γυμνασίου που μπορεί να παρουσιάσει και να αριθμήσει τα ορθογραφικά τους λάθη ώστε να τα μελετήσουμε ως προς αυτές τις μεταβλητές. Βασικός σκοπός επομένως της παρούσας έρευνας, είναι η εστίαση στην ανίχνευση των ορθογραφικών λαθών σε γραπτό λόγο αγοριών και κοριτσιών (ανά φύλο) που φοιτούν και στις τρεις τάξεις του Γυμνασίου (ανά τάξη/ηλικία), χωρίς την απόκλιση γραπτών από μαθητές/τριες με Γ2 την ελληνική και κατ' επέκταση διαφορετική καταγωγή.

Ο θεμελιώδης άξονας της εργασίας λοιπόν διαιρείται σε συγκεκριμένους στόχους. Ειδικότερα i) η καταγραφή και η επεξεργασία των γραπτών σε ειδικό ημιαυτόματο πρόγραμμα για τη δημιουργία του ΜΣΚ σε μορφή txt., ii) ο εντοπισμός και η κατηγοριοποίηση των ορθογραφικών λαθών, iii) η επισημείωση των ορθογραφικών λαθών εντός του ΜΣΚ iv) η καταμέτρηση των λαθών αυτών (ποσοστιαία κατανομή), iv) η σύγκριση τους βάσει κατηγορικών μεταβλητών (φύλο/ ηλικία-τάξη, καταγωγή). Όσον αφορά την καταγωγή, στόχος της έρευνας είναι να μελετήσει αθροιστικά ένα μικρό αριθμό από γραπτά αλλοδαπών μαθητών και μαθητριών του Γυμνασίου που προέκυψαν, ώστε να διερευνηθεί η σχέση των ορθογραφικών λαθών από παιδιά με Γ1 την ελληνική και παιδιά με διαφορετική μητρική γλώσσα. Η εξέταση όλων των προαναφερθέντων, θα οδηγήσει στην εξαγωγή συμπερασμάτων σχετικά με τις κατηγορίες που γίνονται τα συχνότερα ορθογραφικά λάθη από τα αγόρια και τα κορίτσια στο Γυμνάσιο. Κατ' επέκταση αυτό θα φανερώσει τη δυσκολία ή μη που έχουν τα παιδιά στην τήρηση των ορθογραφικών κανόνων. Αυτοί σχετίζονται κυρίως με λάθη στο γραμματικό μόρφωμα, στην κατάσταση της οπτικής τους μνήμης, που τους βοηθάει να διατηρούν αναλλοίωτες στη μακροπρόθεσμη και βραχυπρόθεσμη μνήμη τις σωστά ορθογραφημένες λέξεις και τέλος, στη φωνολογική αδυναμία καταγραφής μια λέξης, που πιθανόν να προκύψει μέσα από τα αποτελέσματα.

6.3. Υποθέσεις της έρευνας

Γενικότερη υπόθεση της έρευνας είναι πως οι μαθητές και οι μαθήτριες στο Γυμνάσιο κάνουν ορθογραφικά λάθη παρά τα έξι χρόνια του Δημοτικού που δίνεται έμφαση στην ανάπτυξη της ορθογραφικής ικανότητας των παιδιών.

Η αρχική αυτή υπόθεση δομείται σε υπό-υποθέσεις που σχετίζονται με τους στόχους της έρευνας και αποτελούν τις ανεξάρτητες μεταβλητές μας:

Ως προς το φύλο:

Υ1α: Τα αγόρια με ελληνική ως Γ1, αθροιστικά κάνουν περισσότερα ορθογραφικά λάθη από τα κορίτσια.

Υ1β: Δεν υπάρχει διαφοροποίηση φύλου ως προς την κατηγορία ορθογραφικών λαθών που γίνονται συχνότερα από τους μαθητές/τριες με Γ1 την ελληνική, δηλαδή αγόρια και κορίτσια κάνουν λάθη στις ίδιες κατηγορίες.

Ως προς την ηλικία/τάξη:

Υ2α: Τα αγόρια (Γ1) σε όλες τις τάξεις (Α', Β', Γ' Γυμνασίου) κάνουν αθροιστικά περισσότερα λάθη από τα κορίτσια.

Υ2β: Όσο μεγαλώνει η τάξη, μειώνονται αριθμητικά τα ορθογραφικά λάθη που γίνονται από τα αγόρια με Γ1 την ελληνική.

Υ2γ: Όσο μεγαλώνει η τάξη, μειώνονται αριθμητικά τα ορθογραφικά λάθη που γίνονται από τα κορίτσια με Γ1 την ελληνική.

Ως προς την καταγωγή:

Υ3α: Οι αλλοδαποί/ές μαθητές/τριες του Γυμνασίου εμφανίζουν τα περισσότερα ορθογραφικά λάθη στις ίδιες ακριβώς κατηγορίες με τους/τις μαθητές/τριες ελληνικής καταγωγής.

6.4. Μορφές ανάλυσης

Η μεθοδολογία που θα χρησιμοποιηθεί, αξιοποιεί τις δυνατότητες και τα πλεονεκτήματα που προσφέρει ο συνδυασμός ποσοτικής και ποιοτικής ανάλυσης των δεδομένων. Η ποσοτική ανάλυση αποβλέπει στο να αναδυθούν οι γενικές κανονικότητες και οι τάσεις που διέπουν κοινωνικά φαινόμενα μέσα από τη διερεύνηση πλήθους περιπτώσεων, εξισορροπώντας έτσι

ιδιομορφίες που χαρακτηρίζουν μεμονωμένες καταστάσεις (Τσιώλης 2011). Κατά τον Leray (2008, 6), η ποιοτική ανάλυση δίνει την ευχέρεια στον/στην ερευνητή/τρια να εξετάσει με μεθοδικότητα τα στοιχεία ενός σώματος στηριζόμενος σε συγκεκριμένες κατηγοριοποιήσεις. Ο συνδυασμός αυτών των μορφών ανάλυσης και η δυνατότητα αλληλοσυμπλήρωσης που παρέχει ευνοεί την αποτελεσματικότητα της έρευνας, αφού κάποιες φορές για την πραγμάτωση μιας ποιοτικής έρευνας αποτελεί προϋπόθεση η περιγραφική ανάλυση, τα αριθμητικά και τα στατιστικά δεδομένα (Δραγώνα 1990).

Η αξιοποίηση της μεθοδολογικής προσέγγισης επιτρέπει να παρουσιαστούν τα δεδομένα με αριθμητικά στοιχεία (ποσοτική ανάλυση), όπου στη συνέχεια τρέπονται σε πληροφορίες που επεξηγούνται μέσω της ποιοτικής ανάλυσης. Αυτή η τεχνική που δομείται στη σύνθεση της στατιστικής διαχείρισης και στην ερμηνεία των ευρημάτων αποτελεί από τη μία, μια ασφαλιστική δικλείδα ανάμεσα στις πιθανότητες μιας απλοποιημένης γενίκευσης αλλά και από την άλλη σε μια γλωσσική υπερ – ανάλυση (Παγκουρέλια και Παπαδοπούλου 2009).

6.5. Μεθοδολογία

Η έρευνα μας όπως ορίζεται από τους στόχους και τις υποθέσεις που τέθηκαν στην αρχή, οδηγούν στην υιοθέτηση συγκεκριμένων μεθόδων ανάλυσης. Πιο συγκεκριμένα:

Την περιγραφική μέθοδο, όπου θα γίνει παρουσίαση των ευρημάτων με βάση τα στατιστικά δεδομένα.

Την αναλυτική μέθοδο, μιας και θα αναλυθούν τα στοιχεία που συλλέχθηκαν.

Την συγκριτική μέθοδο, στην οποία θα στηριχτεί η σύγκριση των δεδομένων.

Την ερμηνευτική μέθοδο, που θα προσπαθήσει συνοπτικά να επεξηγήσει τα δεδομένα που συγκεντρώθηκαν.

6.6. Μαθητικό Σώμα Κειμένων

Γενικότερα, ένα μαθητικό σώμα κειμένων αποτελείται από μια αρκετά μεγάλη και σε υποσύνολα δομημένη συλλογή από αυθεντικά κείμενα. Στη δική μας περίπτωση το δείγμα αποτελείται από εκατόν δεκατέσσερα (114) γραπτά κείμενα παραγωγής λόγου στο μάθημα της Νεοελληνικής γλώσσας και Λογοτεχνίας από μαθητές και μαθήτριες των τριών τάξεων του Γυμνασίου. Συλλέχθηκαν από μη τυπικές μορφές εκπαίδευσης (φροντιστήρια/οικοδιδασκαλεία) αφού εγκρίθηκε η απαιτούμενη άδεια από τους γονείς/κηδεμόνες (βλ. παράρτημα *Έντυπο Γονικής Συγκατάθεσης*). Έγινε προσπάθεια ώστε να

αναδειχθεί η συστηματικότητα της συλλογής των κειμένων με βάση γλωσσολογικά και κοινωνικά κριτήρια προκειμένου να επιτευχθεί μια *εξωτερική ομοιογένεια* στα γραπτά, όπως άλλωστε οφείλει να κάνει ένα αυθεντικό Σώμα Κειμένου. Ο αριθμός των μαθητών και των μαθητριών της έρευνας σίγουρα δεν είναι ιδιαίτερα μεγάλος, ωστόσο το υλικό που συλλέχτηκε κρίνεται ικανοποιητικό για τη δημιουργία του μαθητικού corpus για τη μελέτη των ερευνητικών μας στόχων.

6.7. Κριτήρια επιλογής του δείγματος

Η παρούσα έρευνα μελετάει την ορθογραφική ικανότητα εκατόν δεκατεσσάρων (114) μαθητών και μαθητριών της Α΄, Β΄ και Γ΄ Γυμνασίου στο νησί της Ρόδου, μέσα από την ανάλυση γραπτού λόγου, κατά τις σχολικές χρονιές 2018-2019 και 2019-2020, αφού πρώτα δημιουργηθεί ένα επισημειωμένο μαθητικό σώμα κειμένου με αυτά τα γραπτά μέσω του ημιαυτόματου προγράμματος AntConc3.5.8w. Το τελικό δείγμα ορίστηκε βάσει των κάτωθι κριτηρίων:

Εθνικότητα: Τα υποκείμενα να αποτελούν ένα ρεαλιστικό δείγμα τάξης Γυμνασίου στο νησί, επομένως δεν αποκλείστηκαν γραπτά αλλοδαπών μαθητών/τριών.

Γλώσσα: Η πλειοψηφία των μαθητών/τριών έχει την ελληνική ως μητρική γλώσσα (Γ1), ωστόσο υπάρχουν μαθητές/τριες που γι' αυτούς αποτελεί δεύτερη (Γ2).

Ηλικία: Η ηλικία των μαθητών/τριών κυμαίνεται μεταξύ 12 και 15 ετών, αφού αυτό είναι το ηλικιακό επίπεδο των παιδιών του Γυμνασίου.

Τάξη: Γραπτά από την Α΄, Β΄ και Γ΄ Γυμνασίου.

Φύλο: Τα γραπτά αποτελούνται τόσο από αγόρια όσο και από κορίτσια.

6.8. Διαδικασία δειγματοληψίας

Το δείγμα προέκυψε από τυχαία δειγματοληψία από μαθητές και μαθήτριες των τριών τάξεων του Γυμνασίου που απασχολούνται ενισχυτικά εκτός σχολικού ωραρίου σε μη τυπικούς φορείς ιδιωτικής εκπαίδευσης. Κάθε ένα από τα τυχαία γραπτά που συγκεντρώθηκαν από αγόρια και κορίτσια και χρησιμοποιήθηκαν για να πραγματοποιηθεί η έρευνα, αντιπροσωπεύει ένα μαθητή/τρια. Δε χρησιμοποιήθηκαν δηλαδή γραπτά ίδιου μαθητή/τριας δεύτερη φορά για επεξεργασία. Πρέπει να αναφέρουμε ωστόσο πως γραπτά μαθητών/τριών με διαγνωσμένη δυσλεξία από τα ΚΕΔΔΥ/ΚΔΑΥ, δεν συμπεριλήφθηκαν στην έρευνα μας, γιατί αυτά τα παιδιά αποτελούν ένα άλλο ερευνητικό κομμάτι.

Αφού συλλέχτηκε το δείγμα, έγινε η προσεκτική μεταγραφή του. Φυσικά αφαιρέθηκαν όλα τα προσωπικά δεδομένα των μαθητών/τριών και σημειώθηκε μόνο το φύλο, η τάξη και η καταγωγή που είναι και τα απαραίτητα στοιχεία που μας χρειάζονται για να τους ερευνητικούς μας σκοπούς. Για το λόγο αυτό τα γραπτά οργανώθηκαν σε έξι (6) φακέλους ανά τάξη. Τρεις (3) για τους/τις μαθητές/τριες με μητρική την ελληνική (Α΄, Β΄, Γ΄ Γυμνασίου) και άλλους τρεις (3) για τους/τις μαθητές/τριες με διαφορετική μητρική γλώσσα (Α΄, Β΄, Γ΄ Γυμνασίου).

Στον πρώτο φάκελο που δόθηκε ο τίτλος *AGYMNASIOU* μπήκαν τα γραπτά των αγοριών και των κοριτσιών της Α΄ Γυμνασίου με μητρική γλώσσα την ελληνική. Τα γραπτά των μαθητών/τριών πήραν αύξουσα σειρά και μετονομάστηκαν. Για τα αγόρια δόθηκε η κωδικοποίηση: ABm001, ABm002..... ABm0016. [όπου A: Α΄ Γυμνασίου, B: Boy (=αγόρι), m: mother language (=μητρική γλώσσα), 001..016: τυχαία σειρά κατάταξης γραπτών]. Αντίστοιχα η κωδικοποίηση για τα κορίτσια: AGm001, AGm002....AGm0016 [όπου G: Girl (=κορίτσι)].

Στο δεύτερο φάκελο που δόθηκε ο τίτλος *BGYMNASIOY* μπήκαν τα γραπτά των αγοριών και των κοριτσιών της Β΄ Γυμνασίου με μητρική γλώσσα την ελληνική. Για τα αγόρια δόθηκε η κωδικοποίηση: BBm001, BBm002...BBm016 [όπου B: Β΄ Γυμνασίου, B: Boy (=αγόρι), m: mother language (=μητρική γλώσσα), 001..018: τυχαία σειρά κατάταξης γραπτών]. Αντίστοιχα και τα κορίτσια: BGm001....BGm016.

Στον τρίτο φάκελο που δόθηκε ο τίτλος *CGYMNASIOY* μπήκαν τα γραπτά των αγοριών και των κοριτσιών της Γ΄ Γυμνασίου με μητρική γλώσσα την ελληνική. Αυτή τη φορά δόθηκε η κωδικοποίηση: CBm001, CBm002....CBm016 [όπου C: Γ΄ γυμνασίου, B: Boy (=αγόρι), m: mother language (=μητρική γλώσσα), 001..018: τυχαία σειρά κατάταξης γραπτών]. Ομοίως και για τα κορίτσια : CGm001....CGm0016.

Στον τέταρτο φάκελο που δόθηκε ο τίτλος *AALLODAPOI* ταξινομήθηκαν τα γραπτά των αγοριών και των κοριτσιών της Α΄ Γυμνασίου που έχουν την ελληνική γλώσσα ως δεύτερη. Για τα αγόρια δόθηκε η κωδικοποίηση: ABs001, ABs002, ABs003 [όπου A: Α΄ Γυμνασίου, B: Boy (=αγόρι), s: second language (= δεύτερη γλώσσα), 001..003: τυχαία σειρά κατάταξης γραπτών]. Ομοίως και για τα κορίτσια : AGs001, AGs002, AGs003.

Αυτή η διαδικασία έγινε και με τον πέμπτο φάκελο. Δόθηκε ο τίτλος *BALLODAPOI*, με τα γραπτά των μαθητών/τριών της Β΄ Γυμνασίου που έχουν την ελληνική ως δεύτερη. Για τα

αγόρια δόθηκε η κωδικοποίηση: BBs001 και BBs002 και για τα κορίτσια: BGs001 ...BGs004.

Τέλος, στον έκτο και τελευταίο φάκελο, δόθηκε ο τίτλος *CALLODAPOI*, με τα γραπτά της Γ΄ Γυμνασίου των μαθητών/τριών με δεύτερη γλώσσα την ελληνική. Στα αγόρια υπήρξε η κωδικοποίηση: CGs001... CG003 και στα κορίτσια: CGs001...CGs003

Στη συνέχεια, αφού επιλέχτηκε το υλικό, τα γραπτά μορφοποιήθηκαν και κωδικοποιήθηκαν σε αρχεία .txt, ώστε να μπορούν να γίνουν διαχειρίσιμα ως σώμα κειμένου από το ελεύθερο λογισμικό AntConc 3.5.8w (2009).

Λόγω του ότι η δημιουργία ενός επισημειωμένου σώματος στο πρόγραμμα AntConc είναι μια πιλοτική προσπάθεια, για να επιβεβαιώσουμε την αξιοπιστία των αποτελεσμάτων που θα εξαχθούν, παράλληλα δημιουργήθηκε και μια Κλείδα Καταγραφής των Ορθογραφικών Λαθών.⁹ Εκεί θα καταγράφονται σε κάθετες στήλες τα ορθογραφικά κάθε μαθητή/τριας και σε παράλληλες στήλες οι κατηγορίες των ορθογραφικών λαθών (βλ. [παράρτημα](#)). Στο τέλος, ανά φύλο και τάξη θα υπάρχουν τα αθροιστικά αποτελέσματα των λαθών ανά κατηγορία, που θα λειτουργήσουν στο τέλος ως επαλήθευση των αποτελεσμάτων του ΜΣΚ.

Επομένως, προχωράμε στα στάδια που θα ακολουθήσουμε για τη μελέτη των ορθογραφικών λαθών: 1) εντοπισμός ορθογραφικών λαθών, 2) καταγραφή και ταξινόμηση στην αντίστοιχη κατηγορία μελέτης λαθών στην Κλείδα Καταγραφής 3) επισημείωση των λαθών στο ΜΣΚ 4) καταμέτρηση 5) αποτελέσματα και σύγκριση λαθών ανά φύλο, τάξη, καταγωγή. 6) συμπεράσματα.

Η ανάλυση λαθών και η γενικότερη κατηγοριοποίησή τους αποτελεί τομέα της εφαρμοσμένης γλωσσολογίας και ειδικότερα της γλωσσοδιδασκτικής, που την τελευταία δεκαετία έχει αρχίσει να προσελκύει το ενδιαφέρον και Ελλήνων/ίδων ερευνητών/τριών. Για την κατηγοριοποίηση των λαθών της δικής μας μελέτης αφού εξετάσαμε το σύστημα κατηγοριοποίησης των Κουτσογιάννη και Παπαδοπούλου (2007) καθώς και των Πρωτόπαπα και Σκαλούμπακα (2010) επιλέξαμε την κατηγοριοποίηση των δευτέρων, ως μια πιο πρόσφατη και ολοκληρωμένη προσπάθεια. Έτσι αποφασίσαμε να εστιάσουμε στις τέσσερις από τις έξι βασικές κατηγορίες των Πρωτόπαπα και Σκαλούμπακα (2010). Επομένως η ανάλυση μας θα περιλαμβάνει κατηγορίες: 1) φωνολογικών 2) γραμματικών - μορφολογικών, 3) ετυμολογικών - οπτικών λαθών (ιστορικής ορθογραφίας), 4) και σε μια ακόμη κατηγορία,

⁹ Ακολουθήθηκε ως πρότυπο η Κλείδα Καταγραφής και Ανάλυσης Ορθογραφικών Λαθών των Μουζάκη κ.α (2010, 344-347) και έγιναν οι απαιτούμενες τροποποιήσεις που θα διευκόλυναν την μελέτη μας.

που θα περιλαμβάνει τη λάθος χρήση κεφαλαίων/μικρών γραμμάτων, τις λάθος χωρισμένες λέξεις, και τις λάθος ενωμένες.

ΟΡΘΟΓΡΑΦΙΚΑ ΛΑΘΗ			
1.Φωνολογικά	2.Γραμματικά/μορφολογικά	3.Ετυμολογικά/οπτικά	4 .Κεφαλαία/ μικρά Λάθος χωρισμένες λέξεις Λάθος ενωμένες λέξεις

Πίνακας1. Κατηγοριοποίηση ορθογραφικών λαθών

Σε ένα πρώτο στάδιο εντοπίστηκαν τα ορθογραφικά λάθη στα γραπτά των μαθητών/τριών και καταχωρήθηκαν στις κατηγορίες που είναι προς εξέταση. Υπήρξε προσπάθεια να καταγραφούν τα ορθογραφικά λάθη ανά λάθος και όχι ανά ανορθόγραφη λέξη. Επομένως, μπορεί η ίδια λέξη να έχει καταμετρηθεί παραπάνω φορές, αν εντοπιστούν περισσότερα από ένα λάθη που ανήκουν σε διαφορετική κατηγορία. Για παράδειγμα η λέξη *κουραζόμενοι* (BBm006), έχει καταχωρηθεί στην κατηγορία λαθών που αφορά τη φωνολογική ορθογραφία, αλλά και στην κατηγορία γραμματικού - μορφολογικού λάθους. Επιπλέον, αν και στην ίδια λέξη υπάρχει παραπάνω από ένα λάθος της ίδιας κατηγορίας, πάλι καταχωρείται όσες φορές συναντάται πχ. *σύτηση* (CGm010), *επιρρεάζουν* (CGm003), *τυρρανά* (CBm015). Όταν το λάθος επαναλαμβάνεται πάνω από μία φορές στο ίδιο γραπτό, προσμετρείται μία μόνο φορά στον πίνακα με τα λάθη. Να επισημάνουμε τέλος, πως στην παρούσα εργασία δεν θα γίνει καταμέτρηση τονικών λαθών. Πολλές φορές γίνονται τονικά λάθη και μάλιστα σε περιπτώσεις που η λειτουργία του τόνου είναι καθοριστική για τη σημασία της λέξης πχ. *πότε-ποτέ* (AGm014), άλλες πάλι φορές ο τόνος παραλείπεται ολοκληρωτικά. Παρ' όλα αυτά στο τέλος της μελέτης, θα δοθεί μια γενικότερη εικόνα της τονικότητας των μαθητών/τριών.

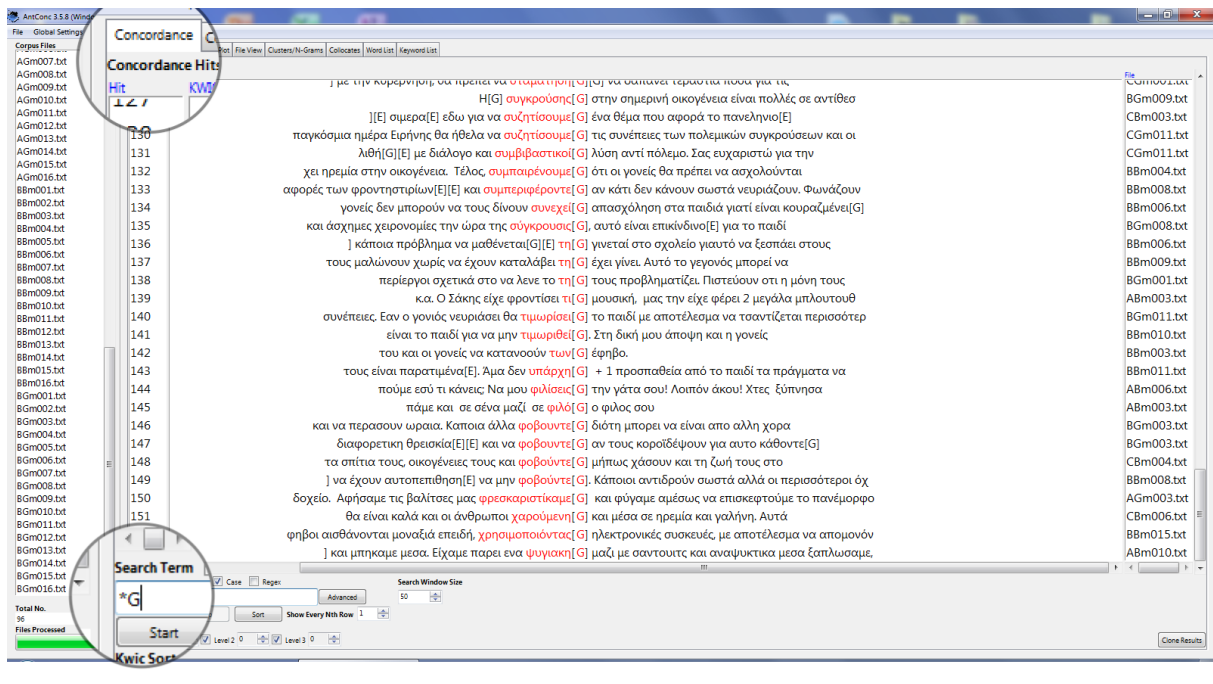
Έπειτα από την καταγραφή και την ταξινόμηση των ορθογραφικών λαθών στην Κλείδα, δημιουργήσαμε ετικέτες - επισημειώσεις των λάθος ορθογραφημένων λέξεων σε κάθε ένα γραπτό/ αρχείο .txt. Πιο αναλυτικά, ορίστηκε ένα γράμμα σε αγκύλη που θα αντιπροσωπεύει την κάθε μια από τις τέσσερις κατηγορίες που έχουμε ορίσει. Όπως: [G] για τα γραμματικά λάθη, [F] για τα φωνολογικά, [E] για τα ετυμολογικά, και [A]¹⁰ για τη λάθος χρήση κεφαλαίων/μικρών, τις λάθος χωρισμένες και τις λάθος ενωμένες λέξεις. Στη συνέχεια, δίπλα

¹⁰ [A] από την κατηγορία 'άλλα λάθη' όπως αναφέρεται στους Πρωτόπαπα και Σκαλούμπακα (2010, 197).

από κάθε λάθος ορθογραφημένη λέξη προστέθηκε σε αγκύλη η κατηγορία του λάθους. Για παράδειγμα οι ενωμένες λέξεις, αυτές δηλαδή που κανονικά αποτελούν σύνολα από δύο λέξεις που ακούγονται μαζί όμως συχνά οι μαθητές/τριες θεωρούν πως είναι μία, κατατάσσονται στην κατηγορία *Αλλα λάθη* [A], έτσι για παράδειγμα η λέξη **μαρέσει** (AGs001), επισημειώθηκε **μαρέσει**[A], αντίστοιχα η λέξη **απέξω** (AGm002) επισημειώθηκε, **απεξω**[A]. Οι λάθος χωρισμένες λέξεις περιλαμβάνουν λέξεις που οι μαθητές/τριες τις χωρίζουν χωρίς να χρειάζεται πχ. και όλας αντί κιόλας (ABm005). Αυτή η επισημείωση έγινε και όλας[A].

Η διαδικασία αυτή ακολουθήθηκε και στα 114 γραπτά. Στην περίπτωση που η μία λέξη είχε πάνω από ένα ορθογραφικά λάθη τότε η μία αγκύλη με το αρχικό της κατηγοριοποίησης που είχαμε ορίσει έμπαινε κολλητά στην άλλη ώστε το πρόγραμμα κατά την αναζήτηση να μπορέσει να μας εμφανίσει όλα τα δεδομένα. Πχ. η λέξη **αποζιμώσης** (GGm005) με δύο ορθογραφικά λάθη, ένα γραμματικό και ένα ετυμολογικό/οπτικό επισημειώθηκε, **αποζιμίωσης**[G][E]. Αυτό συνέβη και στην περίπτωση που μια λέξη εμφάνιζε δύο λάθη στην ίδια κατηγορίας. Πχ. η λέξη **συνυθήσαμε** (AGm003) επισημειώθηκε ως **συνυθήσαμε**[E][E].

Ακολούθησε η αναζήτηση των λαθών από τον πίνακα συχνοτήτων του AntConc, έγινε η συνολική καταμέτρηση και έπειτα η καταμέτρηση ανά τάξη, φύλο και καταγωγή. Εν συνεχεία, δημιουργήθηκαν πίνακες αριθμητικών δεδομένων με τα λάθη σε μορφή ποσοστού επί τοις εκατό (%) σε κάθε κατηγορία.



Εικόνα 1. Στιγμιότυπο οθόνης από την αναζήτηση γραμματικών λαθών στο συμφραστικό πίνακα.

6.9. Περιγραφή τελικού δείγματος

Για τη συλλογή μαθητικού υλικού συμπεριλήφθηκαν όπως προαναφέρθηκε γραπτά από μαθητές και μαθήτριες Γυμνασίου μη τυπικών δομών εκπαίδευσης από το νησί της Ρόδου. Η έρευνα πραγματοποιήθηκε με υλικό των σχολικών χρονιών 2018-2019 και 2019-2020 στο πλαίσιο του μαθήματος της Νεοελληνικής Γλώσσας και της Νεοελληνικής Λογοτεχνίας και πιο συγκεκριμένα κατά την παραγωγή γραπτού λόγου. Τα θέματα αναπαραγωγής σχετίζονταν με αντίστοιχη ενότητα διδασκαλίας, επομένως δεν είναι κοινά πάντα για όλους τους μαθητές/τριες. Παρόλα αυτά η παραγωγή λόγου δε πραγματοποιήθηκε παρουσία της υποφαινόμενης ερευνητριας, αλλά στο πλαίσιο του καθημερινού μαθήματος. Επομένως δεν υπήρξε καμία παρεμβατικότητα, παρά μόνον η φυσική ελεύθερη έκφραση των μαθητών και μαθητριών. Εξάλλου κύριο μέλημα ήταν η αποτύπωση της «αυθεντικής» γραπτής αναπαραγωγής τους. Γι' αυτό το λόγο, θεωρείται πως ο γραπτός λόγος που παρήγαγαν οι μαθητές και οι μαθήτριες, μέσα από την παρατήρηση των ορθογραφικών λαθών δίνει την αντιπροσωπευτική εικόνα τους.

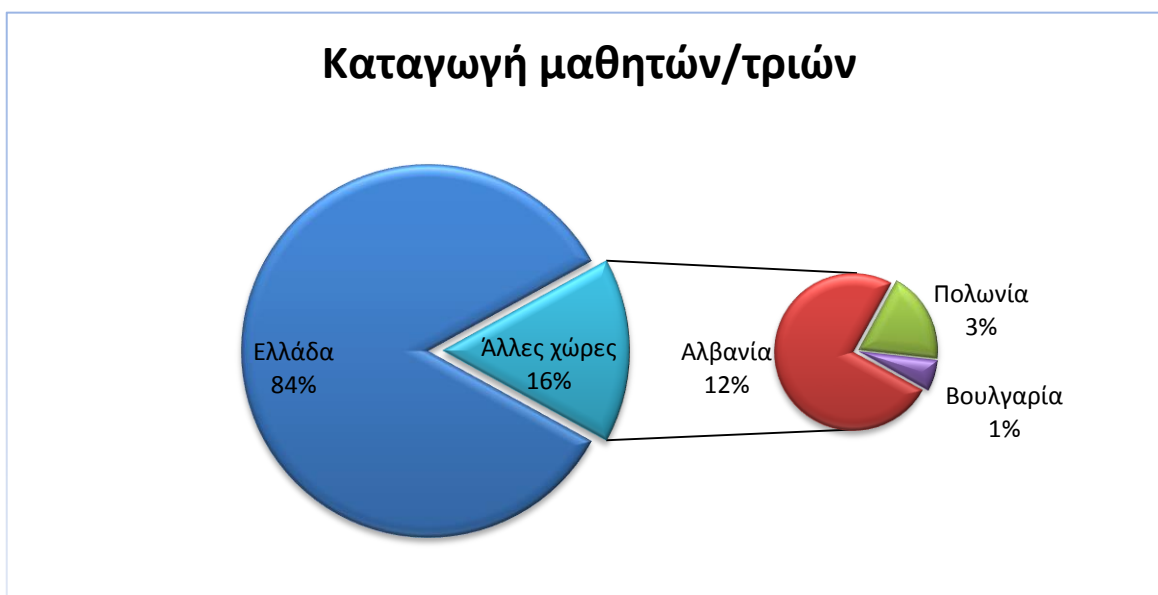
Τα δεδομένα συλλέχθηκαν από ένα σύνολο 114 γραπτών, 96 των φυσικών ομιλητών/τριών της ελληνικής και 18 μη φυσικών. Έγινε προσπάθεια συλλογής ισόποσων γραπτών ανά τάξη

και ανά γλώσσα (συνεπώς και καταγωγή) ώστε να είναι πιο αξιόπιστη η κατανομή των ποσοστών.

	Γ1	Γ2	Σύνολο
Α' Γυμνασίου	32	6	38
Β' Γυμνασίου	32	6	38
Γ' Γυμνασίου	32	6	38
Σύνολο	96	18	114

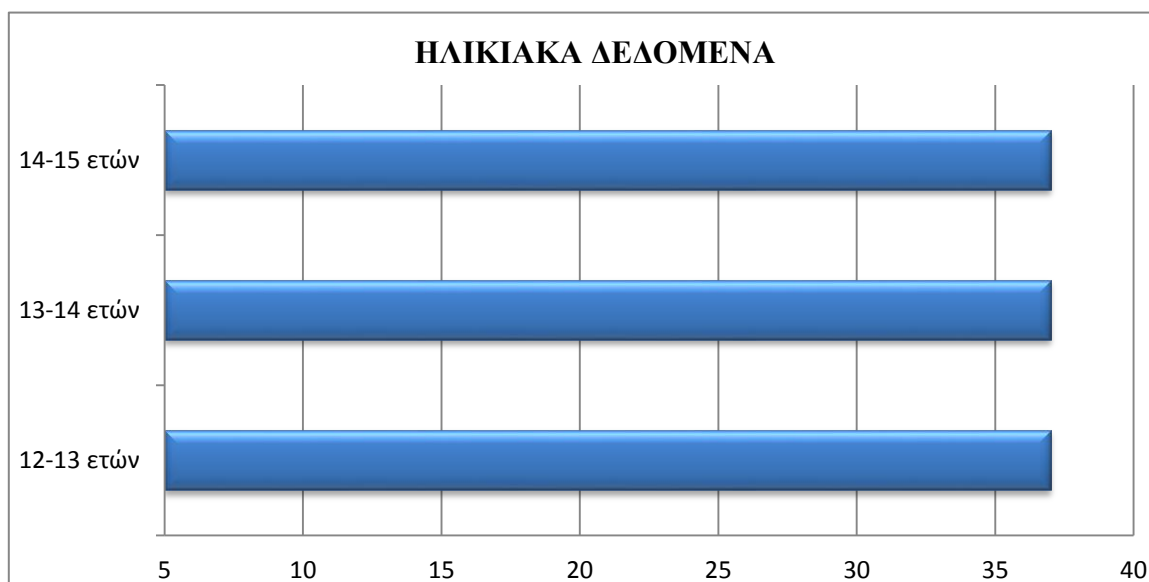
Πίνακας 2. Αριθμός γραπτών ανά τάξη και γλωσσική ομάδα

Παρόλο που δε θα γίνει ανάλυση λαθών ως προς τη χώρα καταγωγής των μαθητών/τριών, η καταγωγή τους απεικονίζεται ακολούθως (γράφημα 1). Αυτό βοηθάει στο να αποκτηθεί μια ολοκληρωμένη εικόνα για την προέλευση των μαθητών/τριών που επικρατεί στα σχολεία της Ρόδου. Όπως φαίνεται και από το γράφημα, από τα 114 γραπτά τα 96 (84%) αποδίδονται σε μαθητές/τριες ελληνικής καταγωγής επομένως τα ελληνικά είναι η μητρική τους γλώσσα (Γ1) και τα 18 (16%) αποδίδονται σε μαθητές/τριες με καταγωγή από άλλες χώρες επομένως είχαν την ελληνική ως δεύτερη γλώσσα (Γ2). Παρατηρούμε πως για τους/τις αλλοδαπούς/ες μαθητές/τριες κυριαρχεί ως μητρική γλώσσα η Αλβανική (12%), ενώ υπάρχουν και μαθητές/τριες με μητρική την Πολωνική (3%) και την Βουλγάρικη γλώσσα (1%).



Γράφημα 1: Κατανομή μαθητών & μαθητριών βάσει καταγωγής

Η ηλικία ήταν ένα από τα βασικά κριτήρια που επιλέχθηκε το δείγμα, αφού οι μαθητές/τριες της έρευνας κυμαίνονται μεταξύ 12 και 15 ετών. Στο ακόλουθο γράφημα παρουσιάζονται τα ηλικιακά δεδομένα:



Γράφημα 2: Κατανομή μαθητών & μαθητριών βάσει ηλικίας.

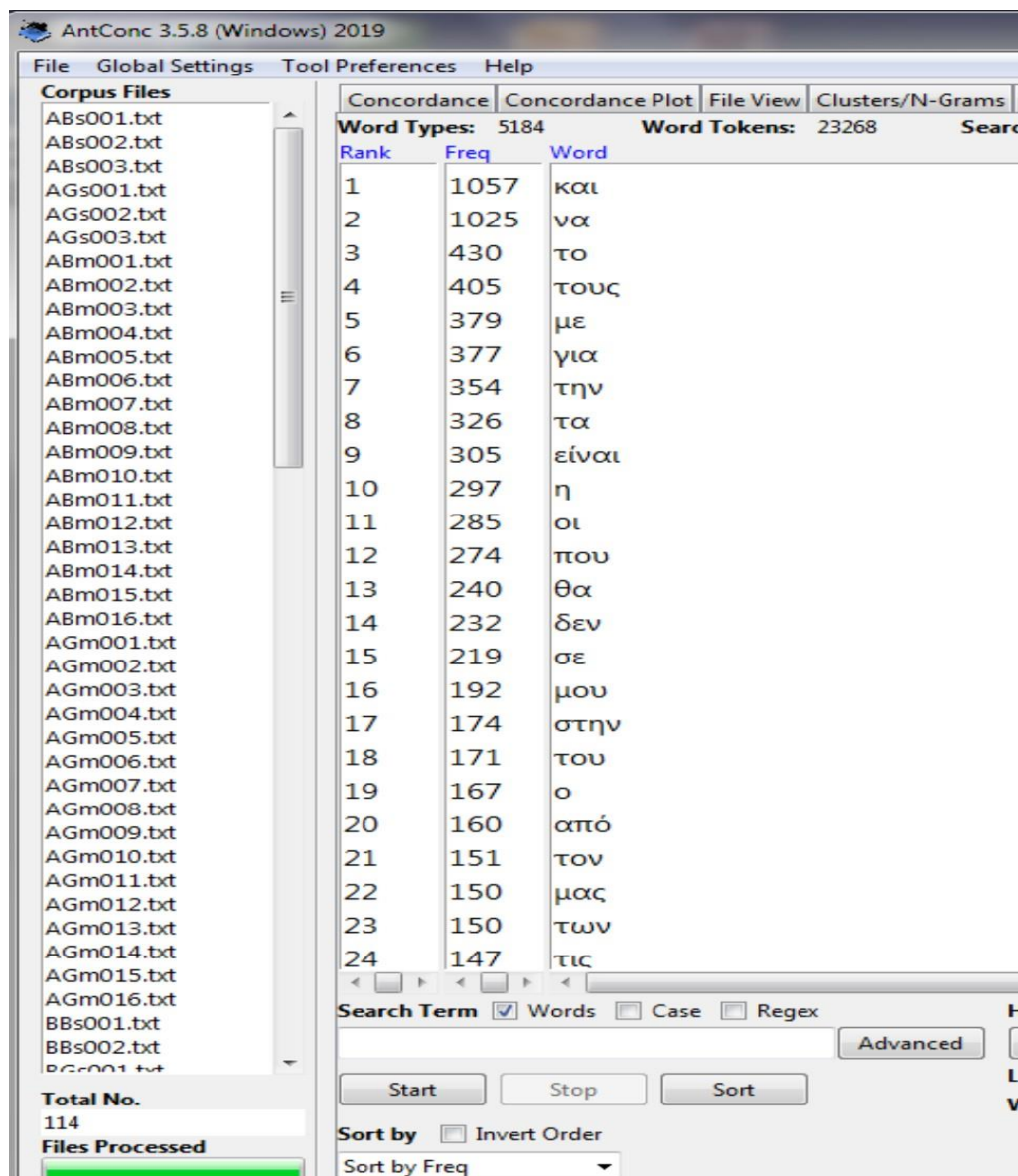
Ως προς το φύλο του δείγματος έγινε προσπάθεια ισόποσης αντιπροσώπευσης, δεδομένου πως υπάρχει πρόθεση για σύγκριση του ποσοστού λαθών ανάμεσα στα δύο φύλα για τους/τις μαθητές/τριες με μητρική την ελληνική. Πιο συγκεκριμένα από το σύνολο των 96 γραπτών από φυσικούς/ες ομιλητές/τριες τα αγόρια καλύπτουν τα 48 γραπτά (50%) και τα κορίτσια τα υπόλοιπα 48 γραπτά του δείγματος (50%).



Γράφημα 2: Κατανομή μαθητών & μαθητριών ανά φύλο

Το μαθητικό σώμα δημιουργήθηκε με τη βοήθεια του λογισμικού AntConc 3.5.8w,(2009) (<http://www.antlab.sci.waseda.ac.jp/software.html>). Επιλέχτηκε αυτό γιατί επιτρέπει ελεύθερη πρόσβαση στο διαδίκτυο, αλλά και γιατί παρέχει συμβατότητα με τα περισσότερα από τα σύγχρονα λειτουργικά συστήματα όπως: Mac, Windows, Linux.

Το ΜΣΚ των 114 γραπτών, αποτελείται από 5184 λεκτικούς τύπους (word types) και από 23268 λεκτικά δείγματα (word tokens). Αναφερόμενοι στους λεκτικούς τύπους εννοούμε τους τύπους εμφάνισης μιας λέξης, ενώ αναφερόμενοι στα λεκτικά δείγματα εννοούμε τον αριθμό των λέξεων που αποτελείται το σώμα.



Εικόνα 4. Στιγμιότυπο οθόνης του AntConc με τον κατάλογο λεκτικών τύπων, λεκτικών δειγμάτων & λεκτικών συχνοτήτων

6.10. Περιορισμοί της έρευνας

Η διαδικασία αντιγραφής των γραπτών για να ταξινομηθούν σε φακέλους και να περαστούν στο ημιαυτόματο πρόγραμμα επεξεργασίας AntConc3.5.8w ώστε να δημιουργηθεί το Μαθητικό Σώμα Κειμένου, σε συνδυασμό με την επισημείωση που ακολούθησε για να μας βοηθήσει στους ερευνητικούς μας σκοπούς, είναι μια χρονοβόρα διαδικασία, που σε τέτοιου είδους έρευνες περιορίζει το μέγεθος του δείγματος.

Η κοινωνικοοικονομική κατάσταση των μαθητών/τριών, το περιβάλλον, το επίπεδο μόρφωσης των γονιών, οι εμπειρίες που είχαν, αποτελούν παράγοντες που μπορεί να επηρεάσουν την ορθογραφική ικανότητα τους, ωστόσο δεν ήταν εφικτό να συνεκτιμηθούν. Επομένως, αναγνωρίζεται ως αδυναμία η εστίαση στη δομική ανάλυση των στοιχείων της πρότασης και μόνον, χωρίς να ληφθεί υπόψη το ευρύτερο εξω-γλωσσικό περιβάλλον.

Επιπλέον, η απουσία επικοινωνίας με τους μαθητές/τριες δεν αφήνει περιθώρια στην εκ βάθους διερεύνηση των γλωσσικών στάσεων και συμπεριφορών όσων συμμετέχουν στην έρευνα. Αυτό θα βοηθούσε ώστε να αποκτήσουμε μια σφαιρική εικόνα για τη διαδικασία εκμάθησης της ορθογραφίας από το Δημοτικό σχολείο, τις δυσκολίες που τυχόν συνάντησαν και τη βαρύτητα που έδωσαν κατά το παρελθόν οι μαθητές/τριες.

Κάτι ακόμα που πρέπει να αναφερθεί είναι πως για τους αλλοδαπούς μαθητές και μαθήτριες έχουμε ελλειπείς πληροφορίες, καθώς δεν γνωρίζουμε το χρονικό διάστημα παραμονής τους στην Ελλάδα. Αυτό θα βοηθούσε να καταλάβουμε μέσα από το γραπτό τους λόγο τον αργό ή ταχύ βαθμό αφομοίωσης τους στο ελληνικό σχολείο. Επιπλέον, ο αριθμός των γραπτών τους που συλλέχτηκε είναι ποσοτικά πολύ μικρότερος από των μαθητών και μαθητριών ελληνικής καταγωγής, κάτι που ωστόσο είναι ρεαλιστικά εφικτό μιας και αυτή η αναλογία επικρατεί στα σχολεία και τα φροντιστήρια της Ρόδου. Παρόλα αυτά δεν πρέπει να αποσιωπήσουμε το γεγονός πως δημιούργησε δυσκολία στο να βγουν συμπεράσματα ποσοτικά τα οποία να είναι με βεβαιότητα αληθή.

Τέλος, πρέπει να τονιστεί η δυσκολία ανάλυσης όλων των ορθογραφικών λαθών με ακρίβεια και αντικειμενικότητα μιας και το θέμα του λάθους είναι σύνθετο και περίπλοκο και δεν είναι ευδιάκριτη πάντα η ταξινόμηση των λέξεων. Τέλος εξαιτίας του μεγάλου εύρους υποκατηγοριών στις βασικές κατηγορίες των ορθογραφικών λαθών και σε συνδυασμό με την πειραματική - πιλοτική δημιουργία του μαθητικού corpus με επισημείωση λαθών, προτιμήθηκε σε αυτή τη φάση η ενασχόληση μόνο με τις βασικές κατηγορίες των ορθογραφικών λαθών.

ΑΠΟΤΕΛΕΣΜΑΤΑ

Αποτυπώνοντας τα αποτελέσματα της καταμέτρησης ως προς τα συνολικά λάθη των μαθητών και μαθητριών της Α', Β' και Γ' Γυμνασίου που έχουν την ελληνική ως μητρική γλώσσα (πίνακας 2.), διαπιστώνεται πως έκαναν στα γραπτά τους 595 ορθογραφικά λάθη των κατηγοριών που θα μελετήσουμε σε σύνολο 20.284 λέξεων (3%), ενώ οι μαθητές και οι μαθήτριες με δεύτερη γλώσσα την ελληνική έκαναν 164 ορθογραφικά λάθη σε σύνολο 2.984 λέξεων (6%).

	ΟΡΘΟΓΡΑΦΙΚΑ ΛΑΘΗ L1	ΠΟΣΟΣΤΟ (%)	ΟΡΘΟΓΡΑΦΙΚΑ ΛΑΘΗ L2	ΠΟΣΟΣΤΟ (%)
Α' Γυμνασίου	221	37%	50	30%
Β' Γυμνασίου	177	30%	65	40%
Γ' Γυμνασίου	197	33%	49	30%
Σύνολο	595	100%	164	100%

Πίνακας 2. Συνολική κατανομή λαθών Α', Β', Γ' Γυμνασίου

Στους επόμενους πίνακες (2α, 2β, 2γ), παρουσιάζεται αναλυτικά ο αριθμός των ορθογραφικών λαθών ανά κατηγορία, φύλο και τάξη, όπως επίσης και η ποσοστιαία στατιστική αποτύπωση τους (%) για τους/τις μαθητές/τριες με μητρική γλώσσα την ελληνική.

Πιο αναλυτικά στην Α' Γυμνασίου, βλέπουμε πως όλοι οι μαθητές και οι μαθήτριες αντιμετωπίζουν μεγαλύτερη δυσκολία στα λάθη που σχετίζονται με την οπτική μνήμη και κατ' επέκταση με την ιστορική ορθογραφία των λέξεων. Το ποσοστό αγγίζει το 68% στα αγόρια ενώ στα κορίτσια φτάνει το 75%. Στην κατηγορία του γραμματικού λάθους, που σχετίζεται με γνώσεις κανόνων γραμματικής, τα ποσοστά λάθους κυμαίνονται ακριβώς στο ίδιο επίπεδο. Σε φωνολογικό επίπεδο τα ποσοστά είναι επίσης πολύ κοντά, με απόκλιση μόνον 1%. Στην κατηγορία άλλα λάθη τα αγόρια παρουσίασαν κατά 6% περισσότερα λάθη, ωστόσο και τα δύο φύλα διατήρησαν χαμηλό ποσοστό.

ΚΑΤΗΓΟΡΙΑ ΟΡΘΟΓΡΑΦΙΚΟΥ ΛΑΘΟΥΣ	ΑΓΟΡΙΑ		ΚΟΡΙΤΣΙΑ	
		Γ1		Γ1
Γραμματικά - μορφολογικά	26	22%	22	22%
Οπτικά - ετυμολογικά	81	68%	76	75%
Φωνολογικά	2	3%	2	2%
Άλλα λάθη	10	8%	2	2%
Σύνολο λαθών	119	100%	102	100%

Πίνακας 2α. Κατανομή ορθογραφικών λαθών Α΄ Γυμνασίου

Ομοίως και στη Β΄ Γυμνασίου, τα ποσοστά ορθογραφικών λαθών που ήταν υψηλότερα εντοπίζονται στις ίδιες κατηγορίες. Με μια πτωτική τάση από τη Α΄ Γυμνασίου, τα λάθη στην κατηγορία ετυμολογικά - οπτικά διατηρούν τα ίδια ποσοστά αγόρια και κορίτσια. Στα γραμματικά - μορφολογικά λάθη τα ποσοστά λαθών είναι πολύ κοντά και στα δύο φύλα. Ακολουθούν με μικρά ποσοστά τα λάθη στις άλλες δύο κατηγορίες.

ΚΑΤΗΓΟΡΙΑ ΟΡΘΟΓΡΑΦΙΚΟΥ ΛΑΘΟΥΣ	ΑΓΟΡΙΑ		ΚΟΡΙΤΣΙΑ	
		Γ1		Γ1
Γραμματικά - μορφολογικά	42	38%	23	35%
Οπτικά - ετυμολογικά	62	56%	37	56%
Φωνολογικά	4	6%	5	8%
Άλλα λάθη	3	3%	1	1%
Σύνολο λαθών	111	100%	66	100%

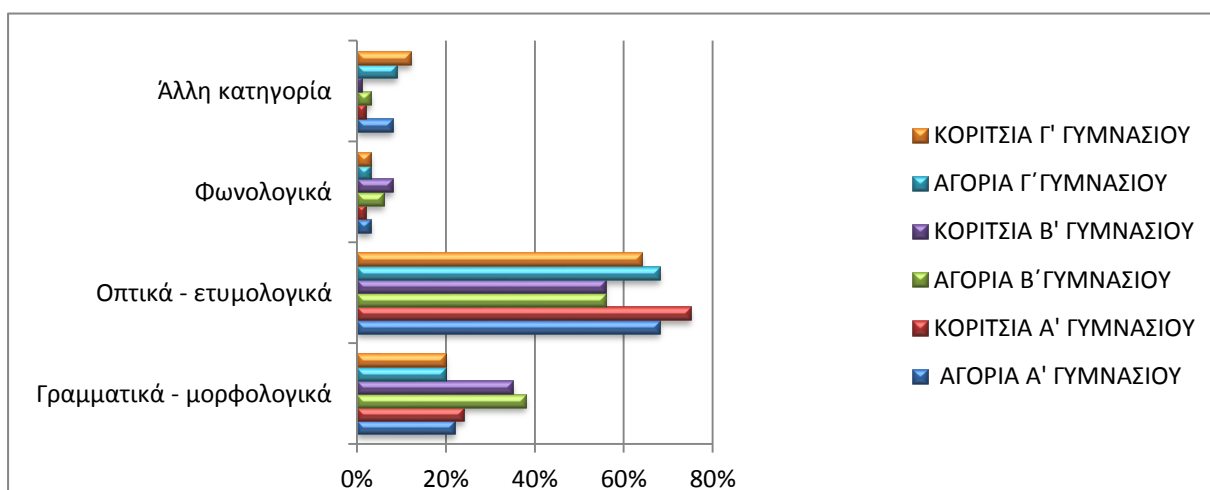
Πίνακας 2β. Κατανομή ορθογραφικών λαθών Β΄ Γυμνασίου.

Στη Γ΄ Γυμνασίου πάλι τα υψηλότερα ποσοστά και για τα δύο φύλα εντοπίζονται στα ετυμολογικά - οπτικά λάθη, με αγόρια και κορίτσια να έχουν μια αυξητική τάση γύρω στο 10% συγκριτικά με τη Β΄ Γυμνασίου. Ακολουθούν τα γραμματικά λάθη, ωστόσο παρατηρήθηκε μια αύξηση του ποσοστού στην κατηγορία «άλλα λάθη» που αφορά κυρίως τη χρήση μικρών/κεφαλαίων και έπονται τα φωνολογικά με μικρότερο ποσοστό (3%).

ΚΑΤΗΓΟΡΙΑ ΟΡΘΟΓΡΑΦΙΚΟΥ ΛΑΘΟΥΣ	ΑΓΟΡΙΑ		ΚΟΡΙΤΣΙΑ	
	Γ1	Γ1	Γ1	Γ1
Γραμματικά - μορφολογικά	20	20%	20	20%
Οπτικά - ετυμολογικά	67	68%	63	64%
Φωνολογικά	3	3%	3	3%
Άλλα λάθη	9	9%	12	12%
Σύνολο λαθών	99	100%	98	100%

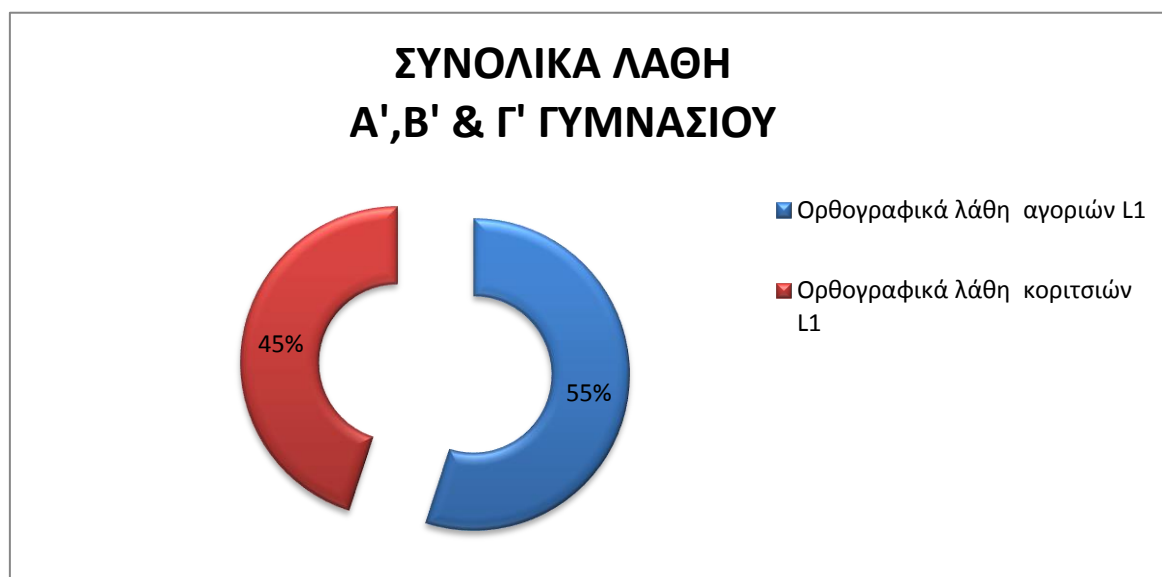
Πίνακας 2γ. Κατανομή ορθογραφικών λαθών Γ΄ Γυμνασίου.

Στο γράφημα 3. που ακολουθεί απεικονίζονται συγκεντρωτικά όλες οι κατηγορίες με τα ορθογραφικά λάθη των μαθητών και μαθητριών με μητρική την ελληνική και στις τρεις τάξεις του Γυμνασίου.



Γράφημα 3. Αναλυτικός πίνακας ορθογραφικών λαθών Α΄,Β,Γ΄ Γυμνασίου μαθητών/τριών με Γ1 την ελληνική.

Στο *Γράφημα 4*, αποτυπώνεται το συνολικό ποσοστό ορθογραφικών λαθών αγοριών και κοριτσιών με μητρική γλώσσα την ελληνική και στις τρεις τάξεις του Γυμνασίου. Το ποσοστό είναι για τα αγόρια 55% (329/595), και για τα κορίτσια 45% (266/595).



Γράφημα 4. Συνολικά λάθη μαθητών & μαθητριών Α',Β' και Γ' Γυμνασίου με Γ1 την ελληνική.

Για να μπορέσουμε να ανακαλύψουμε σε ποιες κατηγορίες κάνουν περισσότερα λάθη οι αλλοδαποί/ες μαθητές/τριες και να τις συγκρίνουμε με τις κατηγορίες που κάνουν λάθη οι φυσικοί ομιλητές/τριες δημιουργήσαμε τους ακόλουθους πίνακες (3α,3β,3γ) που παρουσιάζουν ανά τάξη τα ορθογραφικά τους λάθη και τις κατηγορίες στις οποίες έγιναν. Επειδή ο αριθμός των γραπτών που συλλέχτηκε είναι πολύ περιορισμένος, αποφασίσαμε να μελετήσουμε τα λάθη τους αθροιστικά και για τα δύο φύλα. Υπολογίσαμε ακολούθως το σταθμικό μέσο αγοριών και κοριτσιών με μητρική γλώσσα την ελληνική σε κάθε κατηγορία ορθογραφικών λαθών για να τους συγκρίνουμε με τους μαθητές και τις μαθήτριες με δεύτερη γλώσσα την ελληνική.

Στον *πίνακα 3α*, παρατηρείται πως οι μαθητές/τριες ελληνικής καταγωγής της Α' Γυμνασίου έχουν αυξημένα τα ποσοστά στα λάθη οπτικής μνήμης 71%, ωστόσο ακολουθούν με μικρή διαφορά (9%) τα λάθη των αλλοδαπών μαθητών στην ίδια κατηγορία και έπονται τα γραμματικά λάθη και στις δύο κατηγορίες να είναι σχεδόν ίσα. (22% vs 23%). Μια διαφορά της τάξεως του 10% παρουσιάζεται στην κατηγορία των φωνολογικών λαθών με τους αλλοδαπούς μαθητές/τριες να έχουν το προβάδισμα. Τα λάθη στη χρήση κεφαλαίων/μικρών, λάθος χωρισμένων λέξεων και λάθος ενωμένων παραμένουν σε χαμηλό επίπεδο και στις δύο κατηγορίες.

ΚΑΤΗΓΟΡΙΑ ΟΡΘΟΓΡΑΦΙΚΟΥ ΛΑΘΟΥΣ	ΑΓΟΡΙΑ/ΚΟΡΙΤΣΙΑ		ΑΓΟΡΙΑ/ΚΟΡΙΤΣΙΑ	
	Γ1		Γ2	
Γραμματικά - μορφολογικά	48	22%	12	23%
Οπτικά - ετυμολογικά	157	71%	32	62%
Φωνολογικά	4	2%	6	12%
Άλλα λάθη	12	5%	2	4%
Σύνολο λαθών	221	100%	52	100%

Πίνακας 3α. Κατανομή ορθογραφικών λαθών ως προς την καταγωγή στην Α΄ Γυμνασίου.

Στη Β΄ Γυμνασίου παρατηρείται σημαντική αύξηση των λαθών στους αλλοδαπούς μαθητές/τριες στην κατηγορία των γραμματικών λαθών αγγίζοντας ίδιο ποσοστό με τα λάθη οπτικής μνήμης (45%). Έπονται τα φωνολογικά διατηρώντας σταθερά επίπεδα σε σχέση με τα γραπτά της Α΄ Γυμνασίου (2% απόκλιση). Στους/στις μαθητές/τριες με Γ1 την ελληνική πρώτα στην κατάταξη έρχονται τα λάθη οπτικής μνήμης και ακολουθούν τα γραμματικά. Τα λάθη στην κατηγορία φωνολογικών λαθών και άλλα λάθη διατηρούνται σε χαμηλά επίπεδα.

ΚΑΤΗΓΟΡΙΑ ΟΡΘΟΓΡΑΦΙΚΟΥ ΛΑΘΟΥΣ	ΑΓΟΡΙΑ/ΚΟΡΙΤΣΙΑ		ΑΓΟΡΙΑ/ΚΟΡΙΤΣΙΑ	
	Γ1		Γ2	
Γραμματικά - μορφολογικά	65	37%	32	45%
Οπτικά - ετυμολογικά	99	56%	32	45%
Φωνολογικά	9	5%	7	10%
Άλλα λάθη	4	2%	0	0%
Σύνολο λαθών	177	100%	71	100%

Πίνακας 3β. Κατανομή ορθογραφικών λαθών ως προς την καταγωγή στην Β΄ Γυμνασίου.

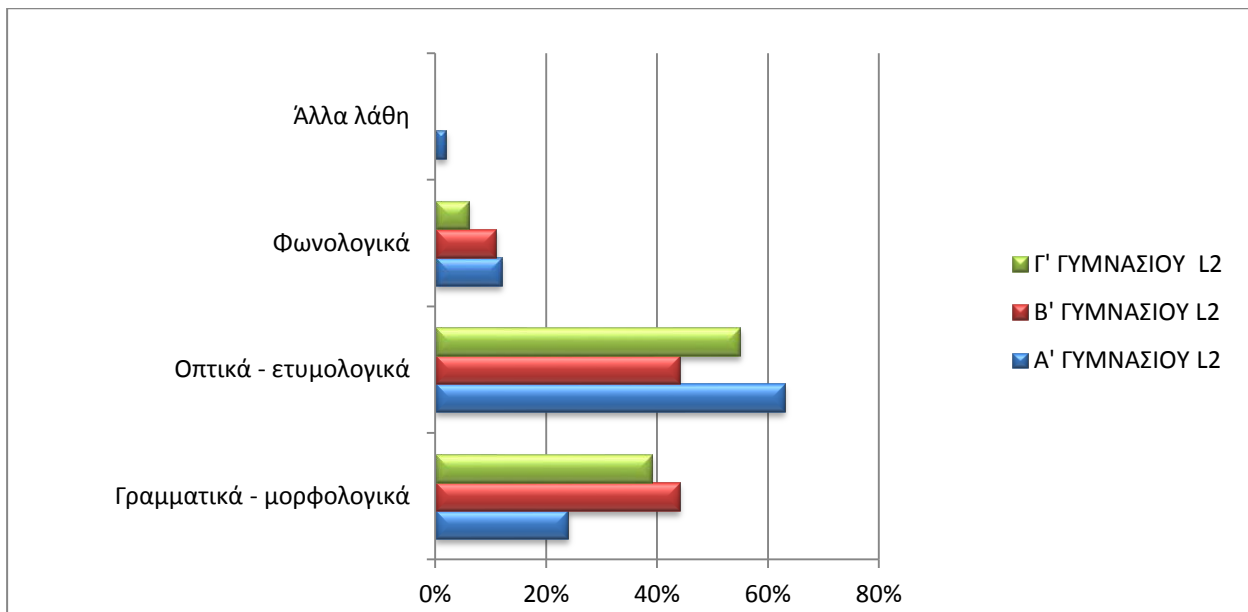
Στη Γ' Γυμνασίου στην πρώτη κατηγορία λαθών είναι τα ετυμολογικά - οπτικά και για τις δύο ομάδες παιδιών και ακολουθούν τα γραμματικά. Σε αυτή την τάξη τα ποσοστά φωνολογικής γραφής συνεχίζουν να παραμένουν σε χαμηλότερα επίπεδα από τις άλλες δύο τάξεις, ωστόσο στους Έλληνες/ίδες μαθητές/τριες υπάρχει μια μικρή αύξηση του ποσοστού που αφορά τη σωστή χρήση κεφαλαίων/μικρών.

ΚΑΤΗΓΟΡΙΑ ΟΡΘΟΓΡΑΦΙΚΟΥ ΛΑΘΟΥΣ	ΑΓΟΡΙΑ/ΚΟΡΙΤΣΙΑ Γ1		ΑΓΟΡΙΑ/ΚΟΡΙΤΣΙΑ Γ2	
	Γραμματικά - μορφολογικά	40	20%	19
Οπτικά - ετυμολογικά	130	65%	27	55%
Φωνολογικά	6	3%	3	6%
Άλλα λάθη	21	11%	0	0%
Σύνολο λαθών	199	100%	49	100%

Πίνακας 3γ. Κατανομή ορθογραφικών λαθών ως προς την καταγωγή στην Γ' Γυμνασίου.

Αυτό που επαληθεύεται και σχετίζεται με τις υποθέσεις της έρευνας μας είναι πως οι αλλοδαποί/ες μαθητές/τριες εμφανίζουν τα περισσότερα λάθη και στις τρεις τάξεις του Γυμνασίου στα λάθη οπτικής μνήμης και γραμματικής, στις ίδιες δηλαδή κατηγορίες που εμφανίζουν τα περισσότερα λάθη και οι Έλληνες/ίδες μαθητές/τριες. Στη φωνολογική κατηγορία τα λάθη δεν είναι ανύπαρκτα ωστόσο συγκριτικά με τις άλλες δύο προαναφερθείσες κατηγορίες και αν λάβουμε υπόψη πως πρόκειται για αλλοδαπούς/ες μαθητές/τριες δεν είναι υψηλά.

Στο γράφημα 5. παρουσιάζονται συγκεντρωτικά και για τις τρεις τάξεις του Γυμνασίου, οι κατηγορίες με τα ορθογραφικά λάθη των αλλοδαπών μαθητών/τριών της ελληνικής γλώσσας.



Γράφημα 5. Αναλυτικός πίνακας ορθογραφικών λαθών Α', Β', Γ' Γυμνασίου μαθητών/τριών με Γ2 την ελληνική.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Η παρούσα έρευνα επιχείρησε τη δημιουργία ενός μαθητικού σώματος κειμένων από γραπτά κυρίως Ελλήνων/Ελληνίδων μαθητών/τριών αλλά και αλλοδαπών. Αφού μελετήθηκαν τα ορθογραφικά λάθη τους, στη συνέχεια ταξινομήθηκαν με βάση την κατηγορία στην οποία ανήκαν και επισημειώθηκαν ώστε να μπορούν να εντοπιστούν και να καταμετρηθούν εύκολα στο πρόγραμμα AntConc3.5.8w. Στη συνέχεια με την καταγραφή των αποτελεσμάτων αναδείχθηκε ο βαθμός κατάρκτησης της ορθογραφικής τους ικανότητας και εντοπίστηκαν οι κατηγορίες ορθογραφικών λαθών με τις μεγαλύτερες αδυναμίες των μαθητών/τριών. Παράλληλα αναλύθηκαν τα αποτελέσματα ώστε να μπορέσουμε να ελέγχουμε τις αρχικές μας υποθέσεις ως προς το φύλο, την ηλικία/τάξη και την καταγωγή των παιδιών. Πιο συγκεκριμένα:

Εξαιτίας του μορφολογικού συστήματος και της ιστορικής ορθογραφίας που έχει η ελληνική γλώσσα, πολλές λέξεις παρατηρήθηκε ότι προκαλούν δυσκολία στους/στις μαθητές/τριες. Κυρίως η δυσκολία αυτή προήλθε από λάθη που αφορούν την οπτική μνήμη των παιδιών. Αν και εμείς για τους σκοπούς της μελέτης μας τα μελετήσαμε ποσοτικά και όχι ως μεμονωμένες

κατηγορίες κατά τους Πρωτόπαπα και Σκαλούμπακα (2010) τα λάθη οπτικής μνήμης κατατάσσονται σε πέντε υποκατηγορίες (βλ. 5.3.2).

Από την ανάλυση ορθογραφικών λαθών στα γραπτά των μαθητών και μαθητριών της Α΄, Β΄ και Γ΄ Γυμνασίου προέκυψε πως η πιο προβληματική κατηγορία ορθογραφικών λαθών τόσο για τα αγόρια όσο και για τα κορίτσια είναι τα λάθη οπτικής μνήμης (ή ιστορικής ορθογραφίας), όπως επιβεβαιώθηκε και από τις Ιγνατιάδου (2008) και Ηλιοπούλου (2009). Πχ. *φενόταν* (AGm004), (οπτικό λάθος σε φωνήεν). Εντοπίστηκαν οπτικά λάθη και στην κατηγορία του θεματικού κανόνα πχ. *νοσοκομίο* (ABm015). Ενδεικτικά, η λέξη οικογένεια, 5/72 γράφτηκε οικογένεια, το ίδιο και η λέξη εφηβεία, 5/9 φορές γράφτηκε (εφηβία). Ομοίως και η χρήση της ερωτηματική αντωνυμία του ουδετέρου *ποιο* αντί του ποσοτικού επιρρήματος *ποιο* (λάθος αναλογίας) πχ. *ποιο πολύ* (ABm009 και BBm011), *ποιο συχνός* (BGm010), *συχνό λάθος που επιβεβαιώνεται και από την Ιγνατιάδου (2008). Πρόβλημα εντοπίστηκε από κάποιους/ες μαθητές/μαθήτριες και στις καταλήξεις άκλιτων λέξεων που δεν εμπίπτουν σε κάποιον μορφολογικό ή φωνολογικό κανόνα (ανώμαλες λέξεις¹¹) πχ. επειδεί* (AGm005), *επίσεις* (AGm005). Τέλος, λάθη εντοπίστηκαν και στην απόδοση φθόγγων με λάθος γράφημα ίδιου φθόγγου πχ. *χάζεβα* (AGm008), *τελαιφταιο* (CBm003), *εφκολα* (BBm012) (οπτικά λάθη σε σύμφωνο).¹² Θα περίμενε κανείς, η αδυναμία αυτή να είναι χαρακτηριστικό στα γραπτά των αλλοδαπών που συναντούν μεγαλύτερη δυσκολία στην απόδοση φθόγγου - γραφήματος, ωστόσο δεν ίσχυσε κάτι τέτοιο στην περίπτωση μας.

Αμέσως μετά ακολουθούν τα γραμματικά - μορφολογικά λάθη. Όσον αφορά το ορθογραφικό λάθος στην κλιτή κατάληξη και τα δύο φύλα συγχέουν το /o/, το /e/ και το /i/, τόσο σε ρήματα όσο και σε ουσιαστικά εξαιτίας της ποικιλίας των καταλήξεων πχ. *σκέφτομε* (ABm012), *είχα αποφασήση* (ABm015), *μαζεύτικα* (ABm014). Στα λάθη αυτά έχει εστιάσει μεγάλο μέρος της γλωσσικής διδασκαλίας και ο χαρακτήρας τους υπερτονίζεται. Μερικά λάθη εντοπίστηκαν και σε άκλιτες καταλήξεις πχ. *χρησιμοποιώντας* (BBm015). Μεγάλη δυσκολία φαίνεται να συναντούν οι μαθητές/τριες με τη χρήση του επιρρήματος πολύ και του επιθέτου πολλές-πολλή-πολύ πχ. *πολύ άστεγοι* (CBm003), *πολύ ησυχία* (ABm008), εύρημα που συνάδει και με προηγούμενες έρευνες (Αναγνωστοπούλου και Κανταρτζή 2008 · Ιγνατιάδου 2008 · Ξάνθη 2017). Τα ορθογραφικά λάθη σε αυτή την κατηγορία είναι λάθη που υπάγονται σε κανόνες

¹¹ Κατά τον Barron (1980, 197, 212) τα λάθη οπτικής μνήμης διακρίνονται σε τρεις κατηγορίες, στα λάθη αναλογίας, στα λάθη ανώμαλων ρημάτων και στα λάθη στο θέμα της λέξης

¹² Σύμφωνα με τον Porrodas (2001), αυτή η κατηγορία λαθών εντάσσεται στα «φωνολογικά λάθη», εννοώντας λάθη που διατηρούν τη φωνολογική ταυτότητα της λέξης. Κατά τους Πρωτόπαπα και Σκαλούμπακα (2010) «φωνολογικά λάθη» είναι τα λάθη που εντοπίζονται στο φωνολογικό επίπεδο και όχι σε εκείνα που το αφήνουν άθικτο.

ορθογράφησης και είναι πιο εύκολο να απομνημονευθούν και να ανακληθούν, σε σχέση με τα οπτικά λάθη. Ωστόσο εξαιτίας της ατελούς εκμάθησης των κανόνων ή της μη συνειδητοποίησης τους, σε συνδυασμό με το ότι έχουν να μελετηθούν οι κανόνες ορθογραφίας από το Δημοτικό έχουν λησμονηθεί · αν προστεθεί και η ελλιπής εξάσκηση, δημιουργούν λάθη στο γραπτό λόγο.

Οι υπόλοιπες κατηγορίες παραμένουν σε χαμηλά επίπεδα. Ελάχιστα λάθη παρατηρήθηκαν σε φωνολογικό επίπεδο πχ. *υοθεσία* (GGs003), (Αντιστροφή δίμηφου φωνήεντος), *Ανέστητη* (BBm008),(αντικατάσταση γραφήματος).

Το ίδιο συνέβη και στην κατηγορία λανθασμένου χωρισμού λέξης με τέσσερις μόνο εμφανίσεις στα 114 γραπτά πχ. *και όλας* (ABm005), πχ. *με τα αντί μετά* (ABm009)και (ABm011), ενώ και οι λάθος ενωμένες λέξεις είχαν ομοίως μόλις τέσσαρις εμφανίσεις στα 114 γραπτά πχ. *απέζω* (AGm002), *εκειπέρα* (AGm005). Δεν έλειψε τέλος, η λάθος χρήση μικρών/κεφαλαίων από μαθητές/τριες και των δύο ομάδων που εξετάστηκαν κυρίως ύστερα από τελεία (έναρξη με μικρό), και ύστερα από κόμμα (έναρξη με κεφαλαίο) καθώς και χρήση μικρού γράμματος σε κύριο όνομα πχ. *άγγελο* (ABm001), *γιώργο* (ABm006), (ABm008)

Λόγω μικρού αριθμού γραπτών από τους/τις αλλόδαπούς/ες μαθητές/τριες καλό θα ήταν να μην γενικεύσουμε τα αποτελέσματα μας, σίγουρα όμως αξίζει να σημειωθούν κάποιες παρατηρήσεις. Και για τους αλλοδαπούς μαθητές/τριες τα υψηλότερα ποσοστά λαθών εντοπίστηκαν στα λάθη οπτικής μνήμης και ακολούθησαν τα γραμματικά λάθη κάτι που επιβεβαιώνεται και στην έρευνα του Αϊδίνη (2006, 39) και της Ηλιοπούλου (2009). Το αποτέλεσμα αυτό δείχνει αναμενόμενο ειδικά για αυτή την κατηγορία μαθητών/τριών γιατί για κάποιον/α που θέλει να μάθει την ελληνική γλώσσα, η ορθογραφία του θέματος μιας λέξης αποτελεί μια ευρεία κατηγορία. Αντίθετα, είναι πρακτικά πιο εύκολο κάποιος να κατακτήσει τη γραμματική, μιας και υπόκειται σε συγκεκριμένους κανόνες και αποτελούν πιο περιορισμένη κατηγορία. Σύμφωνα με τον Varnhagen (1995, 282), οι μαθητευόμενοι έχοντας μικρή μορφολογική γνώση προσπαθούν να χρησιμοποιήσουν στρατηγικές για τις κλινόμενες λέξεις προσέχοντας έτσι ορθογραφικές συμβάσεις που έχουν σχέση με τις καταλήξεις και καταλήγουν να παραμελούν το θέμα. Παρατηρήθηκε λοιπόν, μια ποσοτική κατάταξη λαθών περίπου ίδια τόσο στους φυσικούς όσο και στους μη φυσικούς ομιλητές της ελληνικής. Πέραν από τις δύο αυτές κατηγορίες, στις υπόλοιπες και δη στα φωνολογικά λάθη που κάποιος θα περίμενε υψηλότερα ποσοστά, τα ποσοστά ήταν αρκετά χαμηλά όπως και στην έρευνα της Ηλιοπούλου (2009). Αυτό είναι σαν να προδίδει τη μακροχρόνια παραμονή τους

στην Ελλάδα, τη φοίτηση τους σε ελληνόφωνο σχολείο και κατ' επέκταση την κατάκτηση ή τουλάχιστον την προσπάθεια κατάκτησης του ορθογραφικού συστήματος παράλληλα με τους υπόλοιπους γηγενείς μαθητές/τριες.

Αξίζει να αναφερθεί, παρόλο που δεν ήταν αρχικός στόχος της μελέτης μας αλλά προέκυψε μέσα από την παρατήρηση των γραπτών, πως μια ακόμη προβληματική κατηγορία για τους/τις μαθητές/τριες ανεξαρτήτως τάξης και καταγωγής είναι ο τονισμός. Χωρίς να γίνει καταμέτρηση του φαινομένου, παρατηρήθηκαν γραπτά με παντελή παράλειψη ή λανθασμένου τονισμού. Συχνά τονίζονταν μονοσύλλαβες λέξεις ενώ παρατηρήθηκε διπλός τονισμός σε κάποιες άλλες.

Πρέπει επίσης να τονίσουμε πως στο σύνολο των 114 γραπτών που εξετάστηκαν, υπήρξαν έξι εκθέσεις μαθητών και μαθητριών που δεν παρουσίασαν κανένα ορθογραφικό λάθος, κάτι που υποδηλώνει την ανομοιογένεια των γλωσσικών επιδόσεων των μαθητών/τριών.

Συνοψίζοντας, οι υποθέσεις της έρευνας μας μερικώς επιβεβαιώθηκαν.

Ως προς το φύλο:

Υ1α) Τα αγόρια στη Γ1 αθροιστικά έκαναν περισσότερα λάθη από τα κορίτσια, με ποσοστιαία διαφορά 10% (45% vs 55%). Αποτέλεσμα που συνάδει και με την Τζωρτζάτου (2012)

Υ1β) Δεν υπήρξαν διαφοροποιήσεις φύλου ως προς την κατηγορία ορθογραφικών λαθών που γίνονται συχνότερα. Αποδείχθηκε μέσα από τη μελέτη μας πως η πιο προβληματική ορθογραφία παρουσιάστηκε στα λάθη οπτικής μνήμης (ή ιστορικής ορθογραφίας) τόσο για τα αγόρια όσο και για τα κορίτσια. Ακολούθησαν και για τις δύο ομάδες τα γραμματικά λάθη με ίδια ποσοστά λαθών στις 2/3 τάξεις του Γυμνασίου.

Άρα ως προς το φύλο οι υποθέσεις μας επιβεβαιώθηκαν πλήρως.

Ως προς την ηλικία/τάξη:

Υ2α) Πράγματι τα αγόρια σε κάθε τάξη του Γυμνασίου εμφάνισαν αθροιστικά περισσότερα ορθογραφικά λάθη από κορίτσια.

Υ2β) Όσο μεγαλώνουν τα αγόρια ηλικιακά και ανεβαίνουν τάξη, τα ορθογραφικά τους λάθη μειώνονται σε συνολικό αριθμό (Α' : 119, Β' : 111 , Γ' 99).

Υ2γ) Όσο μεγαλώνουν τα κορίτσια ηλικιακά και ανεβαίνουν τάξη τα ορθογραφικά λάθη δεν ακολουθούν πτωτική τάση και στις τρεις τάξεις του Γυμνασίου. Από την Α΄ στη Β΄ πράγματι υπάρχει μείωση ορθογραφικών και μάλιστα μεγάλη (Α΄: 102, Β΄: 66). Όμως στη Γ΄ Γυμνασίου τα ορθογραφικά αποκτούν αυξητική τάση (από Β΄:66 στη Γ΄: 98). Άρα η μελέτη μας ως προς την κατηγορία ηλικία/ τάξη επαλήθευσε 2/3 υποθέσεις μας.

Ως προς την καταγωγή:

Οι αλλοδαποί μαθητές/ριες του Γυμνασίου που μελετήσαμε κάνουν ορθογραφικά λάθη στις ίδιες κατηγορίες με τους φυσικούς ομιλητές/τριες της ελληνικής. Τα υψηλότερα ποσοστά εμφανίστηκαν στα ορθογραφικά λάθη οπτικής (ή ιστορικής) μνήμης (Α΄ τάξη: 62%, Β΄ τάξη: 45% και Γ΄ τάξη: 55%) και ακολούθησαν τα λάθη στη γραμματική (Α΄ τάξη 23%, Β΄τάξη: 45% και Γ΄ τάξη: 29%).

Όπως φάνηκε και από τους πίνακες η παραγωγή φωνολογικών λαθών διατηρήθηκε σε χαμηλά επίπεδα και στις τρεις τάξεις Γυμνασίου και μάλιστα υπήρχε πτωτική τάση όσο μεγάλωνε η τάξη των μαθητών/τριών. Αυτό υποδεικνύει πως έχουν αποκτήσει σε μεγάλο βαθμό τη φωνολογική επίγνωση που είναι απαραίτητη προϋπόθεση για την κατάκτηση της ορθογραφίας (Διαμαντή 2010, 108).

Ανακεφαλαιώνοντας τα ευρήματα της μελέτης μας, μπορούμε να πούμε πως ως προς το φύλο, τα αγόρια παρουσίασαν ποσοτικά περισσότερα ορθογραφικά λάθη και στις τρεις τάξεις του Γυμνασίου. Η διαφορά αυτή φτάνει το 10% μεταξύ αγοριών και κοριτσιών. Ως εκ τούτου, ο παράγοντας φύλο φαίνεται να επηρεάζει σε ένα μικρό ποσοστό τις γλωσσικές επιδόσεις των μαθητών/τριών. Όσον αφορά την ηλικία/τάξη των μαθητών/τριών δεν προέκυψαν σαφή συμπεράσματα · κατά πόσον όσο μεγαλώνουν ηλικιακά οι μαθητές/τριες κατακτούν σε καλύτερα επίπεδα την ορθογραφική ικανότητα μιας και τα κορίτσια παρουσίασαν αύξηση λαθών από τη Β΄ στη Γ΄ Γυμνασίου. Θα περίμενε κανείς πως όσο έρχονται σε επαφή με τη γλώσσα, εξασκούν την οπτική τους μνήμη και εξοικειώνονται με την μορφή των λέξεων, γι' αυτό και παράγουν λιγότερα λάθη όσο μεγαλώνουν. Ωστόσο αυτό δεν επαληθεύεται στην περίπτωση των κοριτσιών. Αντίθετα αποδεικνύει πως δεν υπάρχει απόλυτη ηλικιακή συσχέτιση και ορθογραφική κατάκτηση σε μαθητές/τριες Γυμνασίου. Τέλος, τα αποτελέσματα που προέκυψαν από τους/τις αλλοδαπούς μαθητές/τριες του Γυμνασίου, οδηγούν στο συμπέρασμα πως οι μαθητές/τριες ανεξαρτήτως καταγωγής κατακτούν τους ορθογραφικούς κανόνες στα ίδια σχεδόν επίπεδα με τους/τις φυσικούς ομιλητές/τριες της ελληνικής. Χωρίς αυτό να αναιρεί πως δεν αντιμετωπίζουν δυσκολίες

κατά τα πρώτα στάδια μάθησης της ελληνικής γλώσσας εξαιτίας της δυσκολίας στη γραφηματική αναπαράσταση του ελληνικού συστήματος που δεν αντιστοιχεί στην προφορά της γλώσσας που προσπαθεί να κατακτήσει. Πρέπει να σημειώσουμε όμως πως πρόκειται για παιδιά Γυμνασίου (όχι Δημοτικού) και δεν γνωρίζουμε πότε ακριβώς ήρθαν στην Ελλάδα. Επομένως, όπως προαναφέρθηκε και άνωθεν, τα χαμηλά ποσοστά σε αυτή την κατηγορία υποδηλώνουν τη χρόνια παραμονή τους στην Ελλάδα, (πιθανότατα να είναι μετανάστες δεύτερης γενιάς) και ως φυσικό επακόλουθο της τριβής τους με τη Γ2 γλώσσα σε όλους τους κοινωνικούς τομείς, δεν αποκλείεται να έχει μετατραπεί και σε Γ1.

Όσον αφορά τη αξιοπιστία του λογισμικού AntConc φάνηκε απόλυτα αξιόπιστο στη δική μας μελέτη. Τα αποτελέσματα μας ταυτίστηκαν ένα προς ένα με τα αθροιστικά αποτελέσματα που προέκυψαν από την Κλείδα Καταγραφής των ορθογραφικών λαθών (ενδεικτικά βλ. [παράρτημα πίνακες 1-4](#)). Επομένως σίγουρα συστήνεται η χρήση του από εκπαιδευτικούς και μαθητές/τριες ως ένα αξιόπιστο εργαλείο για τη διδασκαλία.

ΠΡΟΕΚΤΑΣΕΙΣ

Για τις ανάγκες της διπλωματικής δημιουργήθηκε ένα αυθεντικό Μαθητικό Σώμα Κειμένων, που αποτελεί την πιο σημαντική συνεισφορά της μελέτης μας, γιατί είναι από τα πρώτα μαθητικά σώματα κειμένων που συγκροτείται από γραπτά μαθητών/τριών με πρώτη γλώσσα την ελληνική. Υπάρχουν φυσικά παρόμοιες προσπάθειες αλλά ως επί το πλείστον αφορούν γραπτά που έχουν την ελληνική ως δεύτερη γλώσσα. Μολονότι η μελέτη αυτή δεν είναι εξαντλητική δίνει μια μελλοντική διάσταση στο corpus για εμπλουτισμό και μελλοντική αξιοποίηση. Θα μπορούσε να εμπλουτιστεί με υλικό από μαθητές/μαθήτριες όλων των βαθμίδων και από διαφορετικές περιοχές της Ελλάδας ενισχύοντας με αυτόν τον τρόπο το corpus αλλά ταυτόχρονα δημιουργώντας και χώρο για νέα ερευνητικά ερωτήματα, με πιο επισφαλή συμπεράσματα.

Παρόλα αυτά, τα ευρήματα της παρούσας μελέτης και τα συμπεράσματα στα οποία καταλήξαμε θα μπορούσαν να αποτελέσουν βοηθητικό εργαλείο για τους διδάσκοντες της νέας ελληνικής γλώσσας στα σχολεία τυπικής εκπαίδευσης. Σε σχολεία δηλαδή που συνυπάρχουν εκτός από μαθητές/τριες με μητρική γλώσσα την ελληνική και μια μειοψηφία από μαθητές/τριες που έχει την ελληνική γλώσσα ως δεύτερη. Αυτό θα ήταν βοηθητικό ώστε οι εκπαιδευτικοί να ανακαλύψουν το επίπεδο των μαθητών/τριών τους (στην προκειμένη

περίπτωση για την κατάκτηση του ορθογραφικού ζητήματος και για τις αδυναμίες τους) αλλά και να αξιολογήσουν την παρεχόμενη διδασκαλία τους. Από την άλλη με εστίαση στη λίστα συχνοτήτων (word list) προκύπτουν οι συχνότερες λάθος ορθογραφημένες λέξεις, όπου με περεταίρω ανάλυση θα επιτευχθεί βαθύτερη κατανόηση και στοχευόμενη αντιμετώπιση των αδυναμιών του μαθητή/τριας.

Είναι επιπλέον εφικτό, με υλικό από το ΜΣΚ που σχεδιάστηκε ή και συνδυαστικά με τα ευρήματα της μελέτης, α) να παραχθούν διδακτικές δραστηριότητες β) να χρησιμοποιηθεί ως έχει για τη διδασκαλία τόσο για την αντιμετώπιση ορθογραφικών λαθών, η γλωσσικών λαθών γενικότερα, αλλά και μέσο του πίνακα συχνοτήτων μπορεί να εμπλουτιστεί με νέες επισημειώσεις ώστε να μελετηθούν φαινόμενα γραμματικά και συντακτικά, συχνές γλωσσικές φράσεις μαθητών/τριών κτλ. γ) ο ίδιος ο/η εκπαιδευτικός μπορεί επίσης να κάνει αξιολόγηση της διδασκαλίας του ή της επίδοσης των μαθητών/τριών του/της. Στη διδασκαλία ο/η εκπαιδευτικός, αν λάβει υπόψη του/της τη συχνότητα των ορθογραφικών λαθών ανά κατηγορία, σε συνδυασμό με την προέλευση των μαθητών/τριών που απευθύνεται και τις δυνατότητες ή αδυναμίες που έχουν (ένας/μία εκπαιδευτικός μετά από ένα διάστημα τριβής μαζί τους, πάντα αποκτά μια ολοκληρωμένη ή σχεδόν ολοκληρωμένη εικόνα), θα μπορέσει να ενισχύσει την ορθογραφική τους ικανότητα ακόμη και στο Γυμνάσιο που τα παιδιά έχουν απομακρυνθεί από το μοτίβο κανόνας/ορθογραφία. Πέραν όμως αυτών θα μπορούσε να αξιοποιηθεί και ως υλικό αυτοβελτίωσης των επιδόσεων των ίδιων των παιδιών μέσα από γραπτά της ίδιας τάξης που προκύπτουν καθ' όλη τη διάρκεια του σχολικού έτους.

Μοναδική και απαραίτητη προϋπόθεση όλων όσων προαναφέρθηκαν, να τολμήσουν οι εκπαιδευτικοί να εντάξουν τα Μαθητικά Σώματα Κειμένων στη διδασκαλία τους είτε ενισχυτικά είτε αξιολογικά.

ΒΙΒΛΙΟΓΡΑΦΙΑ

ΕΛΛΗΝΟΓΛΩΣΣΗ

Αϊδίνης, Α. 2006. «Κατάκτηση της ανάγνωσης και της ορθογραφημένης γραφής στις δύο πρώτες τάξεις του Δημοτικού Σχολείου. Δύο συσχετιζόμενες ή ανεξάρτητες διαδικασίες». Στο *Μελέτες για την ελληνική γλώσσα- Πρακτικά της 26^{ης} Ετήσιας Συνάντησης του Τομέα Γλωσσολογίας της Φιλοσοφικής Σχολής Α.Π.Θ.*, 14-15 Μαΐου, 2005, 34-44. Θεσσαλονίκη: Αριστοτέλειο Πανεπιστήμιο.

Αϊδίνης, Α. 2010. «Η ανάπτυξη των στρατηγικών ορθογραφημένης γραφής στα ελληνικά». Στο *Ορθογραφία: μάθηση και διαταραχές*, επιμ. Α. Μουζάκη, και Α. Πρωτόπαπας, 151-164. Αθήνα: Gutenberg.

Αναγνωστόπουλος, Γ. Π. 1964. Ορθογραφία. Ιστορική επιθεώρησης αυτής από των αρχαίων χρόνων μέχρι της σήμερα. Στο *Μεγάλη Ελληνική Εγκυκλοπαίδεια*, 10ος τόμ., κεφ. Γλώσσα, 108-109. Αθήνα: Φοίνιξ.

Αναγνωστοπούλου, Μ., και Ε. Κανταρτζή. 2008. «Τα ορθογραφικά λάθη των μαθητών/τριών του Δημοτικού Σχολείου στη γραπτή έκφραση». Στο *Ελληνική Παιδαγωγική & Εκπαιδευτική Έρευνα Πρακτικά συνεδρίου τόμος Β*, επιμ. Α.Τριλιανός, και Ι. Καραμήνας, 995-1004. Αθήνα: Ατραπός.

Αναστασιάδη-Συμεωνίδη, Α., Ε. Βλέτση, Μ. Μητσιάκη, Β. Μποζονέλος, και Β. Χούμα 2008. «Τα γλωσσικά λάθη μαθητών της ελληνικής ως δεύτερης γλώσσας και ο ρόλος της Γ1 στις πολυπολιτισμικές τάξεις του Γυμνασίου». Στο *Πρακτικά του Διεθνούς Συνεδρίου με θέμα: Ευρωπαϊκό Έτος Διαπολιτισμικού Διαλόγου: Συνομιλώντας με τις γλώσσες πολιτισμού*, 12-14 Δεκεμβρίου, 2008, 597-612. Θεσσαλονίκη: Αριστοτέλειο Πανεπιστήμιο. Ανακτήθηκε 25-04-2020. <https://www.frl.auth.gr/sites/congres/Interventions/GR/anastassiadis-symeonidis.pdf>

Αντωνίου-Κρητικού, Ι., Μ. Γιάγκου, Β. Κάντζου και Σ. Σταμούλη. 2010. «Ο μαθητής ως ερευνητής: αξιοποίηση σωματών κειμένων στη γλωσσική διδασκαλία». Ανακτήθηκε 25-04-20. https://www.academia.edu/18281560/%CE%9F_%CE%BC%CE%B1%CE%B8%CE%B7%CF%84%CE%AE%CF%82_%CF%89%CF%82_%CE%B5%CF%81%CE%B5%CF%85%CE%BD%CE%B7%CF%84%CE%AE%CF%82_%CE%B1%CE%BE%CE%B9%CE%BF%CF%80%CE%BF%CE%AF%CE%B7%CF%83%CE%B7_%CF%83%CF%89%CE%BC%CE%AC%CF%84%CF%89%CE%BD_%CE%BA%CE%B5%CE%B9%CE%BC%CE%AD%CE%BD%CF%89%CE%BD_%CF%83%CF%84%CE%B7_%CE%B3%CE%BB%CF%89%CF%83%CF%83%CE

[%B9%CE%BA%CE%AE %CE%B4%CE%B9%CE%B4%CE%B1%CF%83%CE%BA%CE%B1%CE%BB%CE%AF %CE%B1](#)

Αντωνοπούλου, Ν., Φρ. Βαλετόπουλος, Μ. Καρακύργιου, Μ. Μουμτζή, και Β. Παναγιωτίδου. 2006. *Ανάλυση λαθών της παραγωγής γραπτού λόγου στις εξετάσεις ελληνομάθειας του ΚΕΓ*. Θεσσαλονίκη: Κέντρο Ελληνικής Γλώσσας. Ανακτήθηκε 30-03-2020. http://www.greek-language.gr/greekLang/files/document/errors/errors_study_a.pdf)

Αρχάκης, Α. 2011. *Γλωσσική διδασκαλία και σύσταση των κειμένων*. 2^η έκδ. Αθήνα: Πατάκη.

Γκότοβος, Α. 1992. *Ορθογραφική μάθηση στο δημοτικό (για εμπειρική έρευνα)*. Αθήνα: Guttenberg.

Γούτσος, Δ. 2003. «Σώμα Ελληνικών Κειμένων: Σχεδιασμός και υλοποίηση». Στο *Πρακτικά του 6^{ου} Διεθνούς Συνεδρίου Ελληνικής Γλωσσολογίας, Πανεπιστήμιο Κρήτης, 18-21 Σεπτεμβρίου 2003*. Ανακτήθηκε 22-12-2019. <http://www.philology.uoc.gr/conferences/6thICGL/gr.htm>

Γούτσος, Δ. και Γ. Φραγκάκη. 2015. *Εισαγωγή στη γλωσσολογία σωμάτων κειμένων*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. Ανακτήθηκε 25-04-2020. <https://repository.kallipos.gr/handle/11419/1932>)

Γούτσος, Δ. 2006. Ανάπτυξη λεξιλογίου- Από το βασικό στο προχωρημένο επίπεδο. Στο Δ. Γούτσος, Μ. Σηφιανού & Α. Γεωργακοπούλου, *Η Ελληνική ως ξένη γλώσσα: Από τις λέξεις στα κείμενα*. Αθήνα: Πατάκη. Ανακτήθηκε 20-2-2020. <https://eclass.uoa.gr/modules/document/file.php/PHIL731/%CE%95%CE%B9%CF%83%CE%B1%CE%B3%CF%89%CE%B3%CE%AE%20%CF%83%CF%84%CE%B1%20%CF%83%CF%8E%CE%BC%CE%B1%CF%84%CE%B1%20%CE%BA%CE%B5%CE%B9%CE%BC%CE%AD%CE%BD%CF%89%CE%BD.pdf>)

Διαμαντή, Β. 2010. «Ορθογραφική ανάπτυξη και φωνολογική επίγνωση». Στο *Ορθογραφία: μάθηση και διαταραχές*, επιμ. Α. Μουζάκη, και Α. Πρωτόπαπας, 107-120. Αθήνα: Gutenberg.

Δραγώνα, Θ. 1990. «Συναρθρώσεις ποσοτικών και ποιοτικών προσεγγίσεων ή όταν τα διχαστικά τείχη καταρρέουν», *Δωδώνη*, 1, Γ', σ. 31-40.

Ηλιοπούλου, Κ. Α. 2009. «Η αξιολόγηση των γραπτών γλωσσικών δοκιμασιών των αλλόγλωσσων μαθητών της δευτεροβάθμιας εκπαίδευσης στο μάθημα της Νέας Ελληνικής

Γλώσσας». Διδακτορική Διατριβή, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκη, 2009
Ανακτήθηκε 10-03-2020. <http://thesis.ekt.gr/thesisBookReader/id/26854#page/1/mode/2up>

Θεοδωροπούλου, Μ., και Γ. Παπαναστασίου. 2001. «Το Γλωσσικό Λάθος». Στο επιμ. Α-Φ Χριστίδης, *Εγκυκλοπαιδικός Οδηγός για τη Γλώσσα*, 199-202. Θεσσαλονίκη: Κέντρο Ελληνικής Γλώσσας.

Ιακώβου, Μ., Ι. Βασιλειάδη- Λιναρδάκη, Φ. Βλάχου, Ο. Δήμα, Μ. Καββαδία, Τ. Κατσινά, Μ. Κουτσουμπού, Σ-Ν. Κύτρου, Χ. Κωστάκου, Φ. Παππά, και Σ. Περρέα. 2017. «ΣΕΠΑΜΕ2: Μια καινούρια πηγή αναφοράς για την ελληνική ως Γ2». Στο *Proceedings of the 12th International Conference on Greek Linguistics*, 16-19 Σεπτεβρίου, 2015, 419-432. Βερολίνο: Ρωμοσύνη.

Ιγνατιάδου, Α. 2008. «Ανάλυση ορθογραφικών λαθών: δεδομένα από γραπτά εξετάσεων του Κέντρου Ελληνικής Γλώσσας» Μεταπτυχιακή εργασία, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, 2008. Ανακτήθηκε 20-03-2020. <http://ikee.lib.auth.gr/record/112886/files/GRI-2009-2678.pdf>

Ινστιτούτο Επεξεργασίας Λόγου. 2019. «ΕΘΕΓ: Εθνικός Θησαυρός Ελληνικής Γλώσσας». Ανακτήθηκε 3-12-2019. <http://www.ilsp.gr/el/services-products/langresources/item/73-etheg-resource>

Ινστιτούτο Νεοελληνικών Σπουδών. χ.χ. «Corpus Προφορικού Λόγου». Ανακτήθηκε 5-12-2019. http://www.ins.web.auth.gr/index.php?option=com_content&view=article&id=506&Itemid=165&lang=el

Ιορδανίδου, Α., και Α. Αμπάτη. 2007. «Τα λάθη των αλλόγλωσσων μαθητών και ο σχεδιασμός της διδασκαλίας της ελληνικής ως δεύτερης/ξένης γλώσσας. Ανακοίνωση που παρουσιάστηκε στο Πανελλήνιο Εκπαιδευτικό Συνέδριο». *Τα Λάθη των Μαθητών: Δείκτες αποτελεσματικότητας ή κλειδιά για τη βελτίωση της ποιότητας της εκπαίδευσης*, , 1-2 Νοέμβριου, 2007, 129-138. Αθήνα: Κέντρο Εκπαιδευτικής Έρευνας

Κακριδής- Φερράρι, Μ. 2008. «Ορθογραφικές μεταρρυθμίσεις: τάσεις και αντιστάσεις». Στο επιμ. Μ. Θεοδοροπούλου, *Θέρμη Και Φως. Licht and Warme*. Αφιερωματικός τόμος στη μνήμη του Α-Φ Χριστίδη, 365-383. Θεσσαλονίκη: Κέντρο Ελληνικής Γλώσσας.

Καραντζόλα, Ε. 1999. «Ορθογραφική ρύθμιση στις «ισχυρές» και «ασθενείς» γλώσσες της Ευρωπαϊκής Ένωσης». Στο επιμ. Α.-Φ. Χριστίδης, «*Ισχυρές» και «ασθενείς» γλώσσες στην Ευρωπαϊκή Ένωση: όψεις του γλωσσικού ηγεμονισμού*, 815-824. Θεσσαλονίκη: Κέντρο Ελληνικής Γλώσσας.

Κατσιμαλή, Γ. 2007. *Γλωσσολογία σε εφαρμογή*. Αθήνα: Καρδαμίτσας.

Κουτσογιάννης, Δ., και Δ. Παπαδοπούλου. 2007. «Παραγωγή γραπτού λόγου και κατανόηση ηλεκτρονικού λόγου από εφήβους: Ανάλυση λαθών. Ανακοίνωση που παρουσιάστηκε στο Πανελλήνιο Εκπαιδευτικό Συνέδριο». *Τα Λάθη των Μαθητών: δείκτες αποτελεσματικότητας ή κλειδιά για τη βελτίωση της ποιότητας της εκπαίδευσης*, 1-2 Νοέμβριου, 2007, 172-187. Αθήνα: Κέντρο Εκπαιδευτικής Έρευνας. Ανακτήθηκε 10/05/2020. http://www.kee.gr/attachments/file/praktika/praktika_lathi.pdf

Κουτσογιάννης, Δ. 2008. «Διδακτική αξιοποίηση των ηλεκτρονικών λεξικών και σωμάτων κειμένων: Θεωρητικό πλαίσιο». Ανακτήθηκε 11-01-2020. http://www.greek-language.gr/greekLang/modern_greek/education/cbt/utilization/theory.html

Μαυροειδή, Α. 2011. «Διαφορές στα ορθογραφικά λάθη μεταξύ Δυσλεκτικών και Κανονικών Ενηλίκων». Μεταπτυχιακή εργασία. Πανεπιστήμιο Μακεδονίας. Ανακτήθηκε 03-05-20. <https://dspace.lib.uom.gr/bitstream/2159/14444/1/MavroeidiAthenaMsc2011.pdf>

Μικρός, Γ., και Μ-Ζ Φουντοπούλου. 2002. «Η αξιοποίηση των Ηλεκτρονικών Σωμάτων Κειμένων στη διδασκαλία της Ισπανικής ως ξένης γλώσσας». Ανακτήθηκε 18-01-2020. <http://users.uoa.gr/~gmikros/Pdf/%C7%20%Ε1%ΕΕ%Ε9%ΕF%F0%ΕF%DF%Ε7%F3%Ε7%20%F4%F9%ΕD%20%C7%D3%CA%20%F3%F4%Ε7%20%Ε4%Ε9%Ε4%Ε1%F3%ΕΑ%Ε1%ΕB%DF%Ε1%20%F4%Ε7%F2%20%C9%F3%F0%Ε1%ΕD%Ε9%ΕΑ%DE%F2%20%F9%F2%20%ΕΕ%DD%ΕD%Ε7%F2%20%Ε3%ΕB%FE%F3%F3%Ε1%F2.pdf>

Μουζάκη, Α. Πρωτόπαπας, Γ. Σιδερίδης, και Π. Σίμος. 2010. «Μια διαδικασία για την αξιολόγηση της ορθογραφίας». Στο *Ορθογραφία: μάθηση και διαταραχές*, επιμ. Α. Μουζάκη, και Α. Πρωτόπαπας, 326-347. Αθήνα: Gutenberg.

Ξάνθη, Σ. Β. 2017. «Ποιοτική ανάλυση λαθών μαθητών Δ΄-ΣΤ΄ δημοτικού σε ορθογραφικά έργα υπαγόρευσης και ελεύθερης γραφής». *Έρευνα στην Εκπαίδευση* 6, αρ. 1: 1-17. Ανακτήθηκε 05-04-2020.

https://www.researchgate.net/publication/312207260_Poiotike_analyse_lathon_matheton_D-ST_demotikou_se_orthographika_erga_ypagoreuses_kai_eleutheres_graphes

Παγκουρέλια, Ε., και Μ. Παπαδοπούλου. 2009. «Κριτική Ανάλυση Λόγου – Ποιοτική Ανάλυση Περιεχομένου: μια πρόταση συνδυαστικής αξιοποίησης για την ανίχνευση της ιδεολογίας των σχολικών εγχειριδίων», *Επιστήμες της αγωγής* 4: 75-87. Ανακτήθηκε 15-03-2020. http://mariapapadopoulou.gr/ar8ra_periodika/40.%20Tracing%20ideology%20by%20means%20of%20CDA%20and%20QCA%20in%20textbooks%20for%20literacy.pdf

Πανταζάρα, Α. Β., και Ε. Μάντζαρη. 2015. «Αναζήτηση σε σώματα κειμένων». Ανακτήθηκε 02-02-2020. https://repository.kallipos.gr/bitstream/11419/5969/1/02_chapter_4.pdf

Παπαδοπούλου- Μανταδάκη Σ. 2004. *Η μάθηση της ορθογραφίας*. Αθήνα: Μεταίχμιο

Παπαδοπούλου, Δ., και Α. Τάντος. 2014. «Η Διδακτική Αξιοποίηση των Σωμάτων Κειμένων Μαθητών στην Ελληνική ως Δεύτερη/ ξένη Γλώσσα». Στο *Μελέτες Αφιερωμένες στην Ομότιμη Καθηγήτρια ΑΠΘ Άννα Αναστασιάδη- Συμεωνίδη*, επιμ. Ζ. Γαβριηλίδου, και Α. Ρεβυθιάδου, 246-262. Εκδόσεις: Σαΐτα. Ανακτήθηκε 25-04-2020. http://ikee.lib.auth.gr/record/269166/files/Meletes_A_Anastasiadi.pdf

Παπαναστασίου Γ. 2001. «Γλώσσα και ορθογραφία». Στο *Εγκυκλοπαιδικός Οδηγός για τη Γλώσσα*, επιμ. Α-Φ. Χριστίδης, 189-193. Θεσσαλονίκη: Κέντρο Ελληνικής Γλώσσας.

Παπαναστασίου, Γ. 2008. *Νεοελληνική ορθογραφία: ιστορία, θεωρία, εφαρμογή*. Θεσσαλονίκη: Ινστιτούτο Νεοελληνικών Σπουδών [Ίδρυμα Τριανταφυλλίδη].

Πετρούνιας, Ε. 1984. *Νεοελληνική γραμματική και συγκριτική ανάλυση*. Θεσσαλονίκη: University Studio Press.

Πολίτης, Π. 2008. *Σώματα κειμένων*. Ανακτήθηκε 31-12-2019. http://www.greek-language.gr/greekLang/modern_greek/bibliographies/corpora/ethel.html

Πρωτόπαπας, Α και Χ. Σκαλούμπακας. 2010. «Κατηγοριοποίηση ορθογραφικών λαθών». Στο *Ορθογραφία: μάθηση και διαταραχές*, επιμ. Α. Μουζάκη, και Α. Πρωτόπαπας, 181-195. Αθήνα: Gutenberg.

Σετάτος, Μ. 1991. *Τα γλωσσικά λάθη και η αντιμετώπισή τους*. Φιλολόγος 63: 26-28

Τάντος, Α., Κ. Αλεξανδρή, Ι. Δόση, Κ. Πουλιού, Π. Σαββίδου, και Γ. Φωτιάδου. 2015. «Ανάλυση λαθών στο ελληνικό σώμα κειμένων μαθητών (ΕΣΚΕΙΜΑΘ): Πρώτα ευρήματα». Στο *Μελέτες για την Ελληνική Γλώσσα. Πρακτικά της 35^{ης} Ετήσιας Συνάντησης του Τομέα Γλωσσολογίας του Τμήματος Φιλολογίας της Φιλοσοφικής Σχολής του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης*, 9-10 Μαΐου, 2014, 719-732. Θεσσαλονίκη: Ινστιτούτο Νεοελληνικών Σπουδών.

Τάντος, Α., Σ. Μαρκαντωνάτου, Α. Αναστασιάδη- Συμεωνίδη, Π. Κυριακοπούλου. 2015. *Υπολογιστική Γλωσσολογία*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. Ανακτήθηκε 18-12-2019. <https://repository.kallipos.gr/handle/11419/2205>

Τάντος, Α. 2015. «Σώματα κειμένων και εφαρμογές». Ανακτήθηκε 02-02-2020. https://repository.kallipos.gr/pdfviewer/web/viewer.html?file=/bitstream/11419/2210/3/02_chapter05.pdf

Τζεβελέκου, Μ., Σ. Σταμούλη, Β. Κάντζου, Γ. Παπαγεωργακόπουλος, Β. Χοντρογιάννη, Σ. Βαρλοκόστα, Μ. Ιακώβου, και Β. Λύτρα. 2008. «Η ελληνική ως δεύτερη γλώσσα: Ευρωπαϊκή κλίμακα αξιολόγησης των γλωσσών και κατάταξη σε επίπεδα γλωσσομάθειας». Στο εκδ. Θ. Δράγωνα και Α. Φραγκουδάκη, *Πρόσθεση, όχι αφαίρεση, Πολλαπλασιασμός, όχι διαίρεση: Η μεταρρυθμιστική παρέμβαση στην εκπαίδευση της μειονότητας της Θράκης*, 155-174. Αθήνα: Μεταίχμιο.

Τζιμώκας, Δ. 2010. «Ηλεκτρονικό σώμα κειμένων (ΗΣΚ) εκμάθησης της νέας ελληνικής ως δεύτερης γλώσσας: Προς ένα ερευνητικό και διδακτικό εργαλείο». Στο *Μελέτες για την Ελληνική Γλώσσα. Πρακτικά της 30^{ης} Ετήσιας Συνάντησης του Τομέα της Γλωσσολογίας του Τμήματος Φιλολογίας της Φιλοσοφικής Σχολής του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης*, 2-3 Μαΐου, 2009, 602-616. Θεσσαλονίκη: Ινστιτούτο Νεοελληνικών Σπουδών. Ανακτήθηκε 25-04-20.

http://ins.web.auth.gr/images/MEG_PLIRI/MEG_30_602_616.pdf

Τζιμώκας, Δ. 2018. «Γραμματικά λάθη των μαθητών της νέας ελληνικής ως δεύτερης γλώσσας και η κωδικοποίησή τους βάσει ηλεκτρονικών σωμάτων κειμένων». Διδακτορική διατριβή, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, 2018. Ανακτήθηκε 04-04-2020. <http://thesis.ekt.gr/thesisBookReader/id/44648#page/1/mode/2up>

Τζωρτζάτου, Κ. 2012. «Ορθογραφικά λάθη ελλήνων και αλλοδαπών μαθητών δημοτικού». Διπλωματική εργασία.. Σχολή Ανθρωπιστικών και Κοινωνικών Επιστημών- Τμήμα Φιλολογίας Πατρών, 2012. Ανακτήθηκε 29-03-2020. <https://nemertes.lis.upatras.gr/jspui/bitstream/10889/5965/1/%CE%9F%CF%81%CE%B8%CE%BF%CE%B3%CF%81%CE%B1%CF%86%CE%B9%CE%BA%CE%AC%20%CE%BB%CE%AC%CE%B8%CE%B7%20%CE%B5%CE%BB%CE%BB%CE%AE%CE%BD%CF%89%CE%BD%20%CE%BA%CE%B1%CE%B9%20%CE%B1%CE%BB%CE%BB%CE%BF%CE%B4%CE%B1%CF%80%CF%8E%CE%BD%20%CE%BC%CE%B1%CE%B8%CE%B7%CF%84%CF%8E%CE%BD%20%CE%94%CE%B7%CE%BC%CE%BF%CF%84%CE%B9%CE%BA%CE%BF%CF%8D.pdf>

Τζωρτζάτου, Κ., Α. Αρχάκης, Α. Ιορδανίδου, και Γ. Ξυδόπουλος. 2019. «Διερεύνηση των ορθογραφικών στάσεων στην πρωτοβάθμια εκπαίδευση: Αποτελέσματα πιλοτικής έρευνας1». Στο *Πρώτη γλώσσα & Πολυγλωσσία Εκπαιδευτικές & Κοινωνικοπολιτισμικές Προσεγγίσεις*, Τύρναβος, 2019, 375- 394. Τύρναβος: Πανεπιστήμιο Θεσσαλίας

Τσεσμελή, Σ. Ν. 2009. «Διδακτικές προσεγγίσεις στην εκμάθηση της ορθογραφίας: μια μελέτη περίπτωσης». *Σύγχρονη Εκπαίδευση: Τρίμηνη Επιθεώρηση Εκπαιδευτικών Θεμάτων*, αρ. 156: 166-180.

Τριανταφυλλίδης, Μ., Κ. Καρθαίο (Κλ. Λάκωνα), Θ. Σταύρου, Αχ. Τζάρτζανο, Β. Φάβη, και Ν. Ανδριώτη. [1941]. 1988. *Νεοελληνική γραμματική (της δημοτικής)*. 3^η έκδ. Θεσσαλονίκη: Ινστιτούτο Νεοελληνικών Σπουδών [Ίδρυμα Μανόλη Τριανταφυλλίδη].

Τσοπάνογλου, Α. 2000. *Μεθοδολογία της επιστημονικής έρευνας και εφαρμογές της στην αξιολόγηση της γλωσσικής κατάρτισης*. Θεσσαλονίκη: Ζήτη.

Τσιώλης, Γ. 2011. «Η σχέση ποιοτικής και ποσοτικής έρευνας στις κοινωνικές επιστήμες: από την πολεμική των «παραδειγμάτων» στις συνθετικές προσεγγίσεις». Στο *Οι κοινωνικές επιστήμες στον 21ο αιώνα: Επίμαχα θέματα και προκλήσεις*, επιμ. Μ. Δαφέρμος, Μ. Σαματάς, Μ. Κουκουριτάκης, Σ. Χιωτάκης, 56-84. Αθήνα: Πεδίο. Ανακτήθηκε 15-03-2020. https://www.researchgate.net/publication/283497743_G_Tsiolis_E_schese_poiotikes_kai_posotikes_ereunas_stis_koinonikes_epistemes_apo_ten_polemike_ton_paradeigmaton_stis_syntetikes_prosengiseis_Sto_MDaphermos_MSamatat_MKoukouritakes_SChiotakes_epim_Oi

Φάκου, Α., Σ. Δρακοπούλου, Χ. Σκαλούμπακας, και Α. Πρωτόπαπας. 2010. « Σύγκριση σχετικής συχνότητας ορθογραφικών λαθών μεταξύ μαθητών με και χωρίς δυσλεξία». Στο *Ορθογραφία: μάθηση και διαταραχές*, επιμ. Α. Μουζάκη, και Α. Πρωτόπαπας, 286-304. Αθήνα: Gutenberg.

Φραντζή, Κ. 2012. *Εισαγωγή στην επεξεργασία σωμάτων κειμένων*. Εκδόσεις: Ιων

Χαλισιάνη, Ι. 2009. «Ηλεκτρονικά σώματα κειμένων και γλωσσική διδασκαλία: διεθνείς αναζητήσεις και διαφαινόμενες προοπτικές για την ελληνική γλώσσα». Στο *Πρακτικά Συνεδρίου 2008, Ευρωπαϊκό Έτος Διαπολιτισμικού Διαλόγου: συνομιλώντας με τις γλώσσες-πολιτισμούς*, 12-14 Δεκεμβρίου, 2008, 818-830. Θεσσαλονίκη. Ανακτήθηκε 29-04-20. <https://www.frl.auth.gr/sites/congres/Interventions/GR/xalisiani.pdf>

Χαλισιάνη, Ι. 2010. «Εκδοχές αξιοποίησης των ηλεκτρονικών σωμάτων κειμένων της νέας ελληνικής στη διδασκαλία της μητρικής γλώσσας». Στο επιμ. Κ. Ντίνας, Α. Χατζηπαναγιωτίδη, Α. Βακάλη, Τ. Κωτόπουλος και Α. Στάμου, *Πρακτικά Πανελληνίου Συνεδρίου με διεθνή συμμετοχή: Η Διδασκαλία της Ελληνικής Γλώσσας ως πρώτης/ μητρικής, δεύτερης/ ξένης*, 4-9 Σεπτεμβρίου, 2009. Φλώρινα: Πανεπιστήμιο Δυτικής Μακεδονίας. Ανακτήθηκε 27-03-2020.

<http://linguistics.nured.uowm.gr/Nimfeo2009/praktika/files/down/poster/xalisiani.pdf>

Χαλισιάνη, Ι. 2011. «Τα ηλεκτρονικά σώματα κειμένων ως περιβάλλοντα για τη διδασκαλία της ελληνικής γλώσσας στη δευτεροβάθμια εκπαίδευση». Στο επιμ. Δ. Κουτσογιάννης, Μ. Παυλίδου, και Ι. Χαλισιάνη, *Μελέτη για την αξιοποίηση των ΤΠΕ στη διδασκαλία της*

Νεοελληνικής Γλώσσας στη Δευτεροβάθμια Εκπαίδευση: Γενικό πλαίσιο και ιδιαιτερότητες, 91-126. Θεσσαλονίκη: Κέντρο Ελληνικής Γλώσσας. Ανακτήθηκε 02-04-2020. http://old.greek-language.gr/sites/default/files/digital_school/p3.1.2_glwssa_bthmia.pdf)

Χαραλαμπίδης, Χ. 2001. «Ορθογραφικά προβλήματα της νεοελληνικής». Μέντορας, *Η ελληνική γλώσσα στον 21^ο αιώνα*, ειδικό τεύχος, 185-202 Παιδαγωγικό Ινστιτούτο: Αθήνα.

Χατζηγεωργίου, Ν., Α. Σπηλιωτοπούλου, Α. Βακαλοπούλου, Α. Παπακωστοπούλου, Στ. Πιπερίδης, Μ. Γαβριηλίδου. 2000. «Εθνικός Θησαυρός Ελληνικών Κειμένων: Σώμα Κειμένων της Νέας Ελληνικής στο Διαδίκτυο. Σώμα κειμένων της Νέας Ελληνικής στο Διαδίκτυο», *Μελέτες για την Ελληνική γλώσσα. Πρακτικά της 21^{ης} Ετήσιας Συνάντησης του Τομέα Γλωσσολογίας της Φιλοσοφικής Σχολής του Αριστοτέλειου Πανεπιστημίου 2001*, 12-14 Μαΐου, 2000, 812-821. Θεσσαλονίκη: Εκδόσεις Κυριακίδη. Ανακτήθηκε 03-01-2020. [http://www.ilsp.gr/administrator/components/com_jresearch/files/publications/APTH2000_HNC_corpus_final\(2\).pdf](http://www.ilsp.gr/administrator/components/com_jresearch/files/publications/APTH2000_HNC_corpus_final(2).pdf)

Χατζησαββίδης, Σ., και Α. Χατζησαββίδου. 2009. *Γραμματική της Νέας Ελληνικής Α' Β' και Γ' Γυμνασίου*. Αθήνα: ΟΕΒΔ

ΞΕΝΟΓΛΩΣΣΗ

Andersen, E.S. 1990. *Speaking with style: The sociolinguistic skills of children*. London: Routledge.

Aston G. 2000. 'Corpora and language teaching'. Στο επιμ. L. Burnard. και T. McEney, *Rethinking language pedagogy from a corpus perspective*, 7-18. Hamburg: Peter Lang.

Barron, R.G. 1980. 'Visual and phonological strategies in reading and spelling'. Στο επιμ. U.Frith, *Cognitive Processes in Spelling*, 195-213. Academic Press: London.

Bartsch, S. 2004. *Structural and functional properties of collocations in English. A corpus study of lexical and pragmatic constraints on lexical co-occurrence*. Tuebingen: Narr.

Bennett, G. R. 2010. An introduction to corpus linguistics. Ανακτήθηκε 8-01-2020.
https://pdfs.semanticscholar.org/8f5d/0e54b7238f2ece1e389944f3e361fa90773b.pdf?_ga=2.81704213.329695755.1578516243-1046145282.1578516243

Braun, S. 2005. “From pedagogically relevant corpora to authentic language learning contents. ReCALL“ 17, αρ. 1: 47-64. Ανακτήθηκε 14-04-2020.
<https://pdfs.semanticscholar.org/d930/abd76269ac049cca92d4a70b2e47b3456ddf.pdf>)

Carter, R. 1998. “Orders of reality: CANCODE, communication, and culture“. ELT Journal, 52, αρ. 1: 43-56.

Chambers, A. 2010. “What is data-driven learning?” Στο επιμ. A. O’Keeffe & M. McCarthy, *The Routledge Handbook of Corpus Linguistics*, 345-358. London: Routledge.

Cheng, W. 2010. “What can a corpus tell us about language learning?” Στο επιμ. A. O’Keeffe & M. McCarthy, *The Routledge Handbook of Corpus Linguistics*, 319-332. London: Routledge.

Cheshire J. 1978. “Present tense verbs in Reading English”. Στο επιμ. P. Trudgil, *Sociolinguistic patterns in British English*, 52- 68. London: Arnold.

Clear J. 2000. “Do you believe in Grammar?”. Στο επιμ. L. Burnard και T. McEney, *Rethinking language pedagogy from a corpus perspective*, 19-30. Hamburg: Peter Lang.

Conrad S. 2000. “Will Corpus Linguistics Revolutionize Grammar Teaching in the 21st Century?” *TESOL Quarterly*, 34, αρ. 3 : 548-560.

‘A List of Corpora of Learner English’. χ.χ Corpus4u. Ανακτήθηκε 26-04-2020.
<https://www.corpus4u.org/threads/2987/>

Dulay, H., M. Burt, και S. Krashen. 1982. *Language two*. New York: Oxford University Press.

Farr, F. 2010. “How can corpora be used in teacher education?” Στο *The Routledge Handbook of Corpus Linguistics*, επιμ. A. O’Keeffe και M. McCarthy, 620-632. London: Routledge.

Fitzpatrick, E., και .S. Seegmiller. 2004. “The Montclair electronic language database project”. *Language and Computers*, Οκτώβριος, 223–237. Ανακτήθηκε 03-03-2020. https://www.researchgate.net/publication/233602196_The_Montclair_Electronic_Language_Database_Project

Gavrilidou, M., P. Labropoulou, N. Papakostopoulou, S. Spiliotopoulou, και N. Nassos. 1998. *Greek corpus documentation. Technical report*. (PAROLE LE2-4017/10369, WP2.9-WP. Αθήνα.

Giagkou, M. και I. Antoniou-Kritikou.2008. “Feedback from Integrating Corpus Consultation in Teaching Greek as Mother Tongue at School”. *Proceeding of 8th Teaching and Language Corpora Conference*, Lisbon, 2-6 Ιουλίου, 2008, 386-391. Λισαβόνα: ISLA.

Gillard, P., και A. Gadsby.1998. “Using a learners’ corpus in compiling ELT dictionaries”. Στο επιμ. S. Granger, *Learner English on Computer*, 159-171. London: Longman.

Gilquin, G., και S. Granger. 2010. “How can data-driven learning be used in language teaching”. Στο *The Routledge Handbook of Corpus Linguistics*, επιμ. A. O’Keeffe και M. McCarthy, 359-370. London: Routledge.

Goutsos, D., O. Hatzidaki, και P. King. 1994. “A corpus-based approach to Modern Greek language research and teaching”. Στο επιμ. I. Philippaki-Warburton, K. Nicolaidis & M. Sifianou, *Themes in Greek Linguistics*. Amsterdam/Philadelphia: Benjamins, 507-513.

Goutsos, D. 2010. “The Corpus of Greek Texts: a reference corpus for Modern Greek”. *Corpora*, 5, αρ. 1: 29-44.

Granger, S. 2002. “A bird’s eye-view of learner corpus research”. Στο επιμ. S. Granger, J. Hung και S. Petch-Tyson, *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, 3-33. Amsterdam: John Benjamins.

Granger, S. 2003a. “Error-tagged Learner Corpus and CALL: A Promising Synergy”. *CALICO Journal*, 20, αρ. 3: 465-480.

Granger, S. 2003b. "The International Corpus of Learner English: A new resource for foreign language learning and teaching and second language acquisition research." *TESOL Quarterly*, 37, αρ. 3: 538-546.

Granger, S. 2008. "Learner corpora". Στο *Corpus Linguistics: An International Handbook*, επιμ. A. Lüdeling και M. Kytö, 1: 259-275. Berlin: Walter de Gruyter.

Iakovou, M., O. Dima, T. Katsina, M. Kavvadia, S.N. Kitrou, C. Kostakou, M. Koutsoubou, F. Pappa, S. Perrea, I. Vasileiadi-Linardaki, και F. Vlachou. 2016. "SEPAME2: The first longitudinal corpus for Greek as an L2". *Epic Series in Language and Linguistics*, 1, 191-201.

James, C. 1998. *Errors in Language Learning and Use. Exploring Error Analysis*. London and New York: Longman.

Johns, T. 1991. "From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning". Στο *English Language Research Journal*, επιμ. T. Johns και P. King, 27-45.

Johns, T. 2002. "Data-driven Learning: The Perpetual Challenge". Στο *Teaching and Learning by Doing Corpus Analysis*, επιμ. B. Kettemann και G. Marko, 107-117. Amsterdam: Rodopi.

Kaltenböck, G., και B. Mehlmauer-Larcher. 2005. "Computer corpora and the language classroom: on the potential and limitations of computer corpora in language teaching". *ReCALL*, 17, αρ. 1, 65-84.

Kennedy, Graeme, D. 1998. *An Introduction to corpus Linguistic*. London: Longman.

Kretzschmar, W.A και J. Lee. 1993. "Spatial analysis of linguistic data with GIS functions." *International Journal of Geographical Information Systems*, 7, 541-560.

"L1 και L2 Acquisition Corpora". χ.χ. Ανακτήθηκε 25-02-2020.

http://martinweisser.org/corpora_site/learner_corpora.html

Leech, G. 1997. "Teaching and language corpora: A convergence". Στο *Teaching and Language Corpora*, επιμ. A. Wichmann, S. Fligelstone, T. McEnery και G. Knowles, 1-23. London: Longman.

Leray, C. 2008. *L'analyse de contenu: de la théorie à la pratique, la méthode MorinChartier*. Canada: Presses de l'Université du Québec.

Lewandowska-Tomaszczyk B., και P.J. Melia επιμ. 1997. *PALC 97: Practical applications in language corpora*. Lodz: Lodz University Press.

Long, M. H. 1991. "Focus on form: a design feature in language teaching methodology". Στο *Foreign Language research in cross-cultural perspective*, επιμ. K.de Bot, R.Ginsberg και C. Kramersch, 39-52. Amsterdam: John Benjamins.

Mair, C. 2002. "Empowering Non-Native Speakers: The Hidden Surplus Value of Corpora in Continental English Departments". Στο *Teaching and Learning by Doing Corpus Analysis*, επιμ. B. Kettemann και G. Marko, 119-130. Amsterdam : Rodopi.

McCarthy, M. 2008. "Accessing and interpreting corpus information in the teacher education context". *Language Teaching*, 41, αρ. 4: 563-574.

McEnery, T. 2005. "Corpus linguistics". *The Oxford handbook of computational linguistics*, 448-463. Ανακτήθηκε 05-12-2019. <http://web.stanford.edu/~eckert/PDF/McEneryWilson.pdf>

Mukherjee, J., και J-M. Rohrbach. 2006. "Rethinking Applied Corpus Linguistics from a language-pedagogical perspective: New departures in learner corpus research". Στο επιμ. B. Kettemann και G. Marko, *Planning, Gluing and Painting Corpora: Inside the Applied Corpus Linguist's Workshop*, 205-232. Frankfurt am Main: Peter Lang.

O'Keeffe, A., και F. Farr. 2003. "Using language corpora in initial teacher education: Pedagogic issues and practical applications". *TESOL Quarterly*, 37, αρ. 3: 389-418.

O’Keeffe, A., M. McCarthy, και R. Carter. 2010. *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press.

PAROLE 1995. “Design and composition of reusable harmonized written language reference corpora for European languages”. *Technical report*. PAROLE Consortium. MLAP : 63-386, WP 4 – Task 1.1.)

Pavlidou, Th.-S. 2012. “The corpus of Spoken Greek: Goals, Challenges, perspectives” . Στο *LREC Proceedings, Workshop 18 (Best Practices for Speech Corpora in Linguistic Research)*, 23-28. Ανακτήθηκε 10-01-2020. http://ins.web.auth.gr/images/M_images/pdf/Pavlidou_Corpus%20of%20Spoken%20Greek.pdf

Porpodas, C.D. 2001. “Cognitive Processes in first grade reading and spelling of Greek, *Ψυχαγωγία*, 8, 384-400.

Roemer, U. 2004. “Comparing real and ideal language learner input: The use of an EFL textbook corpus in Corpus Linguistics and language teaching”. Στο *Corpora and language learner*, επιμ. G. Aston, S. Bernardini και D. Stewart, 151-168. Amsterdam: John Benjamins.

Römer, U. 2008. “Corpora and language teaching”. Στο *Corpus Linguistics: An International Handbook*, επιμ. A. Lüdeling και M. Kytö., 1, 112-131. Berlin: Walter de Gruyter.

Rundell, M. και P. Stock. 1992. “The corpus revolution”, *English Today*, 30,9-14 | 31, 21-32| 32,45-51

Rundell, M. 2008. “he corpus revolution revisited”. *English Today*, 24, αρ. 01 ,23-27.

Sampson, R. 2001. *Empirical Linguistics*. London: Continuum

Sinclair, J. 1996. *EAGLES Preliminary Recommendation on Text Typology*. Διαθέσιμο στο <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.28.1988&rep=rep1&type=pdf>

Sinclair, J. 1997. “Corpus Evidence in Language Description”. Στο *Teaching and Language Corpora*, επιμ. A. Wichmann, S. Fligelstone, T. McEnery και G. Knowles, 27-32. London: Longman.

Sripicharn, P. 2010. “How can we prepare learners for using language corpora?” Στο *The Routledge Handbook of Corpus Linguistics*, επιμ. A. O’Keeffe και M. McCarthy, 371-384. London: Routledge.

Strassel, S., και C. Cieri. 2001. “Data and Annotations for Sociolinguistics: A Corpus-Based Approach to Sociolinguistic Research”. Στο *Penn Linguistics Colloquium*, 3-4 Μαρτίου, 2001. Φιλαδέλφεια. Ανακτήθηκε 25-04-2020. <https://www ldc.upenn.edu/sites/www ldc.upenn.edu/files/plc2001-data-annotations-sociolinguistics.pdf>

Tantos, A. (υπό δημ.). “The Greek Discourse Relations’ Corpus and its Multiple Discriminant Analysis Classifier.” Στο *Aspects of Corpus Linguistics: Principles, implementations, challenges*, Research Center for Scientific Terms and Neologisms: Academy of Athens.

Tantos, A., G. Chatziioannidis, K. Lykou, M. Papatheohari, A. Samara, και K. Vlachos. 2017. “Corpus C58 and a data-analytic approach to the interface between intra- and inter-sentential linguistic information”. Στο *Proceedings of 12th International Conference on Greek Linguistics*. Chair for Modern Greek Studies and the Center for Modern Greece (CeMoG), 961-976. Βερολίνο: Πανεπιστήμιο Freie.

Tono, Y. 2003. “Learner corpora: design, development and applications”. Στο *Proceedings of Corpus Linguistics* επιμ. D. Archer, P. Rayson, A. Wilson, and A. McEnery. Lancaster: Lancaster University, 800-809. Ανακτήθηκε 30-03-2020. <http://ucrel.lancs.ac.uk/publications/cl2003/papers/tono.pdf>

Trudgill P. 1974. *The Social Differentiation of English in Norwich*. Λονδίνο: Cambridge University Press.

Tsalidis, Ch., G. Orphanos, E. Mantzari, M. Pantazara, Ch. Diolis, και A. Vagelatos. 2007. “Developing a Greek Biomedical Corpus for Text Mining”, *Corpus Linguistics Conference*,

27-30 Ιουλίου 2007. Ανακτήθηκε 10-01-2020.
<http://www.birmingham.ac.uk/research/activity/corpus/publications/conference-archives/2007-birmingham.aspx>

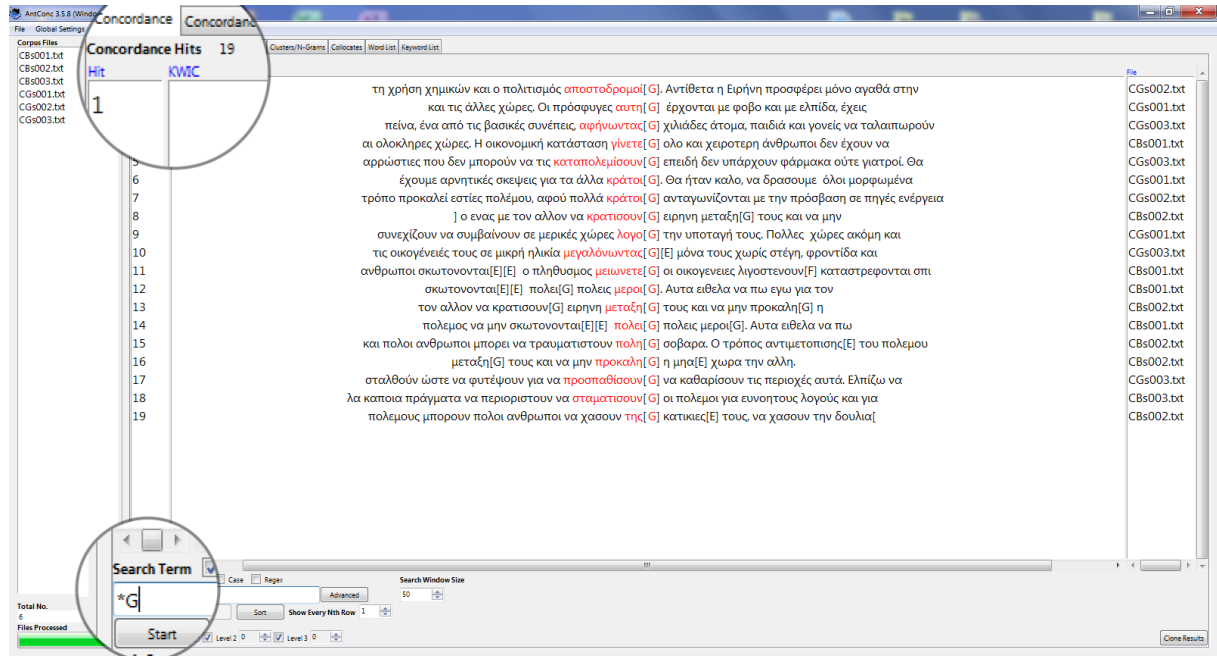
Vaughan, E. 2010. “How can teachers use a corpus for their own research?” Στο *The Routledge Handbook of Corpus Linguistics*, επιμ. A. O’Keeffe και M. McCarthy, 471-487. London: Routledge.

Vienna-Oxford International Corpus of English. 2003. Ανακτήθηκε 27-03-2020.
<https://translate.google.com/translate?depth=1&hl=el&prev=search&rurl=translate.google.gr&sl=en&sp=nmt4&u=https://www.univie.ac.at/voice/>

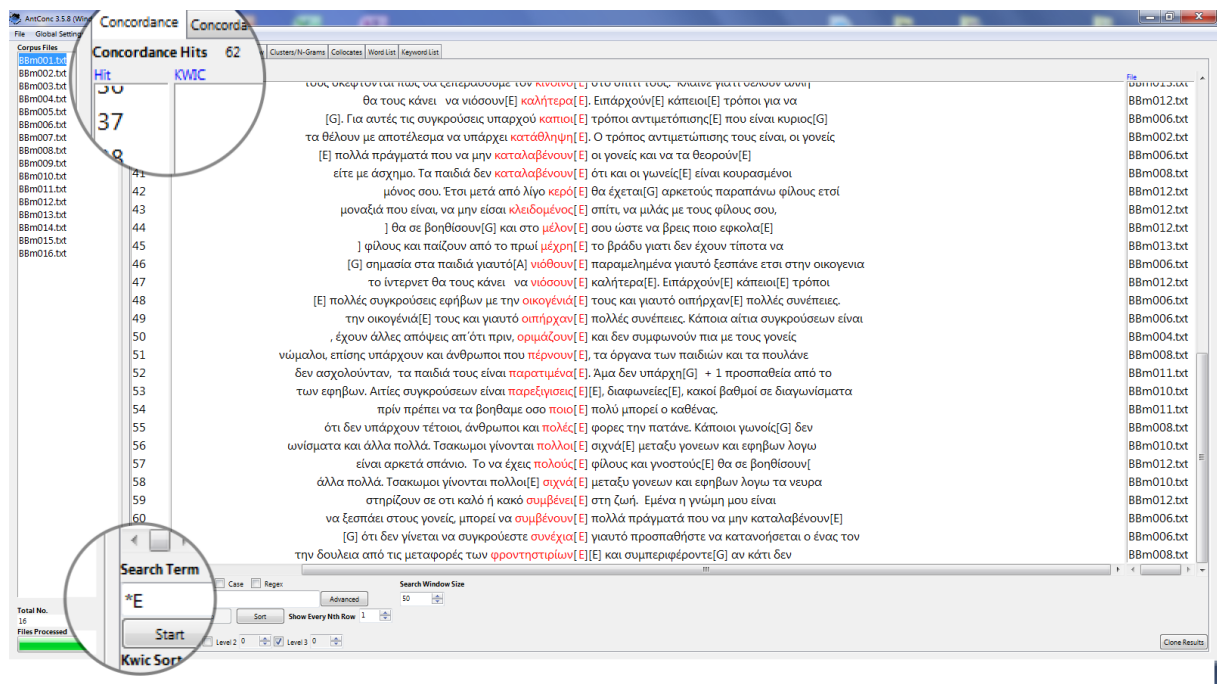
Wichmann, A., S. Fligelstone, T. McEnery, και G. Knowles. 2007. *Teaching and Language Corpora*. London: Longman

ΠΑΡΑΡΤΗΜΑ

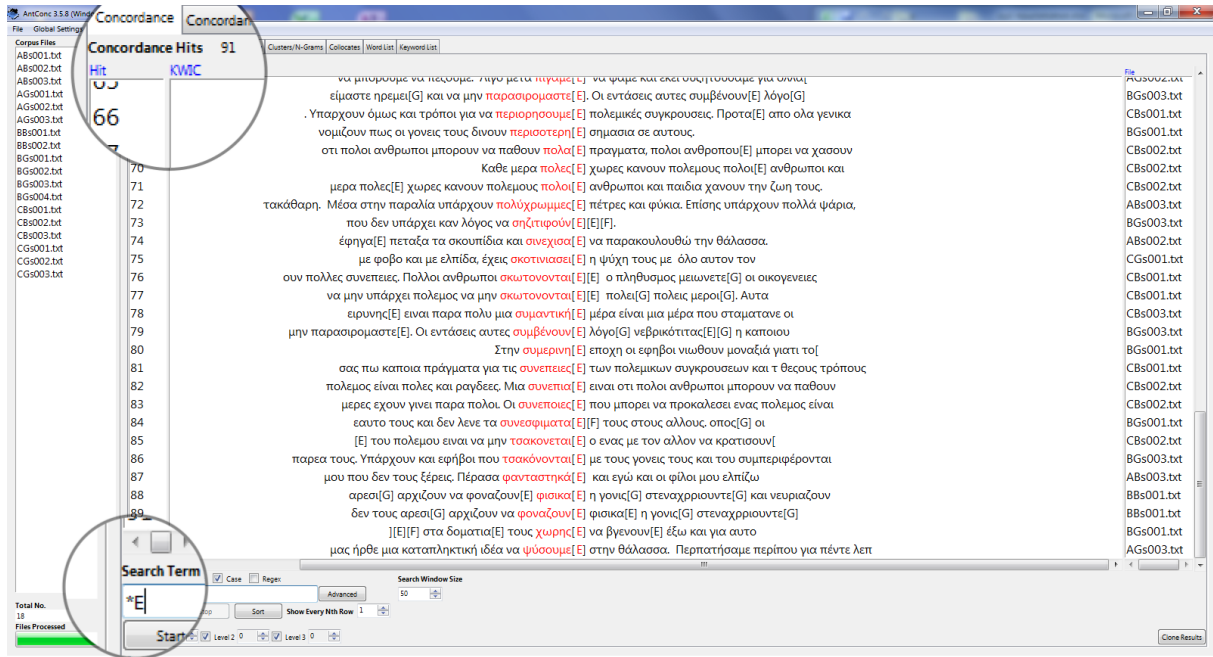
Ενδεικτικές εικόνες από τον συμφραστικό πίνακα του ΜΣΚ στο πρόγραμμα AntConc κατά την καταμέτρηση των ορθογραφικών λαθών



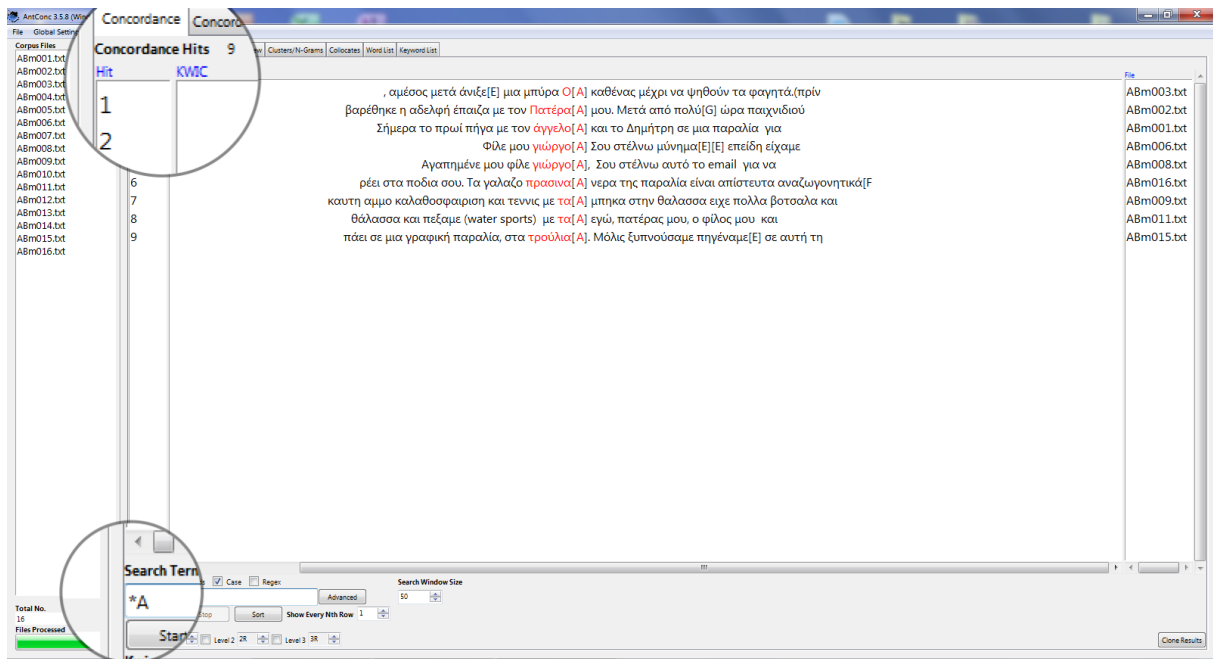
Πίνακας 1. Στιγμιότυπο από τον φάκελο CGYMNASIOY/ALLODAPOI κατά την αναζήτηση γραμματικών λαθών



Πίνακας 2. Στιγμιότυπο από τον φάκελο BGYMNASIOY/BOYS κατά την αναζήτηση ετυμολογικών λαθών



Πίνακας 3. Στιγμιότυπο από τον φάκελο ALLODAPOI/A/B/C GYMNASIOY κατά την αναζήτηση ετυμολογικών λαθών



Πίνακας 4. Στιγμιότυπο από τον φάκελο A GYMNASIOY/BOYS κατά την αναζήτηση «Άλλα λάθη»

ΚΛΕΙΔΑ ΚΑΤΑΓΡΑΦΗΣ ΚΑΙ ΑΝΑΛΥΣΗΣ ΟΡΘΟΓΡΑΦΙΚΩΝ ΛΑΘΩΝ

Λέξη / Γραπτό	γραμματικά [G]	ετυμολογικά [E]	φωνολογικά [F]	άλλο [A]	Εμφάνιση
AGm001	0				
AGm002					
τελίωσα		1			1
απέξω				1	1
κρυστάλινα		1			1
φενόταν		1			1
συνυθήσαμε		2			2
AGm003					
κρητικές		1			1
φρεσκαριστίκαμε	1				1
πείρα		1			1
καΐκα		1			1
αμουδιά		1			1
κρυσταλένια		3			3
AGm004					
μεγάλο (αντί μεγάλη)	1				1
πανλευκά			1		1
φενόταν		1			1
πολύ (αντί πολλοί)	1				1
εξηπυρετικοί		2			2
αναροτιέσαι		1			1
απεριώριστη		1			1
AGm005					
βρείσκομε	1	1			2
μπένεις		1			1
ριχά		1			1
φίνικες		1			1
πολύ (αντί πολλή)	1				1
ισιχεία		3			3
βαριέμε	1				1
ηλιοθεραπία		1			1
καβαλίσω	1				1
επίσεις		2			2
πήγενα		1			1
παραλίες (αντί παραλίας)	1				1
επειδεί		1			1
βαρέθικαμε	1				1
εννώητε	1	2			3
αμμέσος	1	1			2
μπίκαμε		1			1
κολυμπίσαμε	1				1
κολοτούμπες		1			1

πολούς		1		1
κολιμπίσαμε	1	1		2
έφηγα		1		1
εκειτέρα			1	1
ξαναζίσω	1			1
τετια		1		1
εμπιρία		1		1
AGm006				
εξαιρετικά		1		1
ξαπλόστρες		1		1
είσουν		1		1
διαδίκτιο		1		1
AGm007				
κύμματα		1		1
λαμπήριζε		1		1
ψυλά		1		1
AGm008				
σήγουρα		1		1
ακριβός	1			1
μοίραζαν		1		1
ζαλήξεσε	1	1		2
μηλούσαμε		1		1
ιδιοφυές (αντί ιδιοφυή)	1			1
AGm009				
κίματα		1		1
γυροπολεία		1		1
πολί (αντί πολύ)	1			1
AGm010				
συνιθησμένο		2		2
κιτρινοπό		1		1
βράγγια (αντί βράγια)			1	1
AGm011				
κύμμα		1		1
AGm012				
λέγετε (αντί λέγεται)	1			1
κρυστάλινα		1		1
όμους		1		1
εμπυρία		1		1
απόχρωση		1		1
AGm013				
φοίνηκες		1		1
σαλιγγάρια		1		1
διάφωρα		1		1
ριχά		1		1
φώτηζε		1		1
γλώσσες		1		1
AGm014				

κρυστάλινο		1			1
κυπαρρίσια		2			2
επισκευθεί		1			1
AGm015					
κύμματα		1			1
βατίρα		1			1
γίνετε	1				1
έστυσε		1			1
κολόνες		1			1
παρακολουθούσαμε	1				1
περίφανος		1			1
AGm016					
κατευθίαν		1			1
πολύ (αντί πολλή)	1				1
ΣΥΝΟΛΟ ΛΑΘΩΝ	22	76	2	2	102

Λέξη / Γραπτό	γραμματικά [G]	ετυμολογικά [E]	φωνολογικά [F]	άλλο [A]	Εμφάνιση
ABm001					
άγγελο				1	1
κατασκηνόσουμε		1			1
λεοφορείο		1			1
πολή		1			1
ABm002					
Πατέρα				1	1
πολύ (αντί πολλή)	1				1
μεσιμέρι		1			1
ABm003					
εκήγει	1	1			2
αμάξη	1				1
αμέσος	1				1
φουσκοτό		1			1
πάνο	1				1
χτιπάει		1			1
τι (αντί τη)	1				1
άνιξε		1			1
Ο αντί ο				1	1
μεσημεριάση	1				1
νυχτόσει		1			1
βήντεο		1			1
βραδιάση	1				1
φήγαμε		1			1
φιλό	1				1
ABm004					
	0				
ABm005					
επικύνδινο		2			2

χτιπήσουμε		1		1
πολλή αντί πολύ	1			1
και όλας			1	1
κουράστηκες (αντί κουραστικές)		1		1
εκκλισάκι		1		1
απλός αντί απλώς	1			1
ABm006				
γιώργο			1	1
μύνημα		2		2
φιλίσεις	1			1
γενούν		1		1
έβγεναν		1		1
κόνους		1		1
κοχείλια		1		1
μπίκαμε		1		1
ABm007				
Γιόργο		1		1
αποφασήσαμε		1		1
θάλασα		2		2
βρησκόμαστε		1		1
χρεισιμοποιουσαμε		1		1
υπιο		1		1
πολλύ (αντί πολύ)		1		1
ABm008				
γιώργο			1	1
αλλωτε		1		1
περναμε		1		1
δραστηριότητες		1		1
επειδεί		1		1
επισεις		1		1
ABm009				
καταπλικτικα		1		1
με τα αντί μετά			1	1
ποιο αντί πιο		1		1
ευτηχως		1		1
ABm010				
κρυστάλινο		1		1
μεσιμέρι		1		1
κολύμβισης		1		1
φουκοτο αντί φουσκωτό		1		1
ψυγιάκη	1			1
προτινε		1		1
ανεουμε		1		1
θαλασα		1		1
μησω		1		1
ABm011				
ξαπλόστρες		1		1

πολύ αντί πολλή	1			1
βηθό		1		1
πέζαμε		1		1
με τα αντί μετά			1	1
κατάδηση		1		1
ABm012				
πολύ αντί πολλή	1			1
κείματα		2		2
σύνεφο		1		1
κατάδησης		1		1
γύρο	1			1
κατάδισης		1		1
υπύρχαν		1		1
εμπιρία		1		1
σκέφτομε	1			1
ABm013				
διηγιθώ	1			1
χιθήκαμε		1		1
μεσημέρη	1			1
παρείγγειλα		1		1
ανυπομονισία	1			1
ABm014				
έκεγε		1		1
παραμήθι		1		1
έκπληκτη αντί έκπληκτοι	1			1
ABm015				
αποφασήση	1	1		2
τρούλια αντί Τρούλια			1	1
πηγέναμε		1		1
φίνικες		1		1
πολλή αντί πολύ (πολλή μεγάλη)	1			1
πολλή αντί πολύ (πολλή καθαρή)	1			1
επιπροσθέτος	1			1
λέγαμαι	1			1
νοσοκομίο		1		1
άχηνος		1		1
πέζαμε		1		1
φουσκοτά		1		1
ελευθεροθεί		1		1
σφηνόσει		1		1
λεοφορίο		1		1
γύρησε		1		1
ορραίες		2		2
ABm016				
ξαπλόστρες		1		1
γαλαζο πράσινα			1	1

αναζωογονητικά			1		1
μίνεις		1			1
επίσει αντί επίσης		1	1		2
κατάδηση		1			1
φοβησμένη		1			1
σπλιά		1			1
ΣΥΝΟΛΟ ΛΑΘΩΝ	26	81	2	10	119

Λέξη / Γραπτό	γραμματικά [G]	ετυμολογικά [E]	φωνολογικά [F]	άλλο [A]	Εμφάνιση
BGm001					
η (αντί οι)	1				1
τη (αντί τι)	1				1
άνχος			1		1
θεωρόντας	1				1
BGm002					
κοινωνικοποιημένα		1	1		2
συγγενικά			1		1
οφελεί		1			1
BGm003					
σινέβη	1	1			2
οικογένια		1			1
κλισμένα		1			1
παρέαις	1				1
θρεισκία		2			2
φοβούντε	1				1
κάθοντε	1				1
κλισμένα		1			1
παραπάνο	1				1
BGm004					
ανηπομονησία		1			1
αντιθέτως (αντί Α)				1	1
κηνιτό		2			2
BGm005					
λόγο (αντί λόγω)	1				1
πετυχένουν		1			1
εφηβίας		1			1
καταλαβαίνουμαι	1				1
BGm006					
επιρεάζει		1			1
λόγο (αντί λόγω)	1				1
παράπωνα		1			1
επιδώσεις		1			1
πρωτεραιότητες		1			1
μεταβήβαση		1			1
BGm007					

αντιμετοπίζονται		1			1
BGm008					
αναφορά (αντί αφορά)			1		1
λόγο (αντί λόγω)	1				1
γωνιός		1			1
ανχώνεται			1		1
συγκρούσις	1				1
επικίνδυνο		1			1
προβληματίζει		1			1
BGm009					
η (αντί οι)	1				1
συγκρούσης	1				1
ασύμαντες		1			1
πολύ (αντί πολλή)	1				1
κλίνεται		1			1
δημιουργούντε	1				1
BGm010					
ποιο (αντί πιο)		1			1
τσακομού		1			1
πολύ (αντί πολλή)	1				1
υπόγλυπα		2			2
αντιμετώπισεις	1				1
λόγο (αντί λόγω)	1				1
BGm011					
πιθανών		1			1
καταλαβαίνεται	1				1
υποτειθέμενη		1			1
ορμόνες		1			1
οικογένιας		1			1
τιμωρίσει	1				1
BGm012					
λόγο (αντί λόγω)		1			1
BGm013					
κοινείς	1				1
προμελετιμένα		1			1
επίσεις		1			1
BGm014					
έλειψη		1			1
μειωνεκτήματα		1			1
BGm015					
	0				
BGm016					
	0				
ΣΥΝΟΛΟ ΛΑΘΩΝ	23	37	5	1	66
Λέξη / Γραπτό	γραμματικά	ετυμολογικά	φονολογικά	άλλο	Εμφάνιση
	[G]	[E]	[F]	[A]	

BBm001				
σύνχρονοι			1	1
BBm002				
διχόνοια		1		1
αυτών (αντί αυτών)	1			1
κατάθληψη		1		1
BBm003				
εφοίβων		1		1
έφηβω	1			1
οι(αντί η)	1			1
αναπτύξη	1			1
εντάσης	1			1
των (αντί τον)	1			1
BBm004				
αλάζουν		1		1
οριμάζουν		1		1
συμπαιρένουμε	1			1
BBm005				
εξωμαληθούν			1	1
BBm006				
γύνονται		1		1
οικογένιά		1		1
οιπήρχαν		1		1
λόγο	1			1
πολλοί (αντί πολλή)	1			1
γιαυτό			1	1
νιόθουν		1		1
συνεχεί	1			1
κουραζμένει	1		1	2
αντιμετόπισης		1		1
κυρίος	1			1
περιστρέφοντε	1			1
καπιοι		1		1
αφστηρεί	1	1		2
ηδικά		1		1
λογο (αντί λόγω)	1			1
εφιβίας		2		2
λήσεται	1	1		2
μαθένεται	1	1		2
τη (αντί τι)	1			1
συμβένουν		1		1
καταλαβένουν		1		1
θεορούν		1		1
κατανοήσεται	1			1
συνέχια		1		1
BBm007				
επηρεάζετε (αντί επηρεάζεται)	1			1

επιεικής (αντί εκεικείς)	1		1
BBm008			
αιτίες		1	1
καταλαβένουν		1	1
γωνείς		1	1
φροντηστιρίων		2	2
συμπεριφέροντε	1		1
πέρνουν		1	1
αρνιτικά		1	1
πολές		1	1
γονοίς	1		1
ενδιαφέροντε	1		1
αφείνουν		1	1
γίνοντε	1		1
κακή (αντί κακοί)	1		1
ανέστητη (αντί αναίσθητοι)	1		2
άχρησται	1		1
ανήκανε	1	1	2
αυτοπεπίθηση		1	1
φοβούντε	1		1
BBm009			
τη (αντί τι)	1		1
εμπιστοσήνη		1	1
ανωσιές		2	2
BBm010			
παρεξιγίσεις		2	2
διαφωνείες		1	1
πολλοί (αντί πολύ)		1	1
σιχνά		1	1
πολλή (αντί πολύ)	1		1
τιμωριθεί	1		1
BBm011			
γονειών		1	1
παρατιμένα		1	1
υπάρχη	1		1
διαλλείμα		2	2
ποιο		1	1
BBm012			
γιαυτό			1
αντικοινωνικοί		1	1
νιόσουν		1	1
καλήτερα		1	1
ειπάρχουν		1	1
κάπειοι		1	1
κλειδομένος		1	1
κερό		1	1
έχεται	1		1

πολούς		1			1	
γνοστούς		1			1	
βοηθίσουν	1				1	
μέλον		1			1	
εφκολα		1			1	
συμβένει		1			1	
BB013						
νοιάζοντε	1				1	
κίνδινο		1			1	
κάποιει	1				1	
ικανεί	1				1	
διαδικτυακούς		1			1	
μέχρη		1			1	
BBm014						
εφηβία		1			1	
πολύ παίζουν				1	1	
BBm015						
χρησιμοποιώντας	1				1	
απομονώνονται		1			1	
λόγο	1				1	
μιλώντας	1				1	
BBm016						
	0					
ΣΥΝΟΛΟ ΛΑΘΩΝ		42	62	4	3	111

Λέξη / Γραπτό	γραμματικά [G]	ετυμολογικά [E]	φωνολογικά [F]	άλλο [A]	Εμφάνιση
CBm001					
Διευθύντρια (αντί δ)				1	1
αλλειωθεί		1			1
συμαθητής		1			1
παγκόσμιας				1	1
ημέρας				1	1
CBm002					
δυσεκατομύρια		2			2
συνέπια		1			1
CBm003					
μαζεφτίκαμε	1	1			2
σιμερα		1			1
συζητίσουμε	1				1
πανεληνιο		1			1
συνεποιές		1			1
αρχιγία		1			1
βρησκεταί		1			1
τελιωσεί		1			1
ητέ		1			1
επανερθη	1		1		2

καταστασει	1		1
πινα		1	1
συνθυκές		1	1
τροφη		1	1
τελαιφαταίο		1	1
αθλεια		1	1
δωθούν		1	1
ανεξαρτιτός		2	2
οραίο		1	1
βρησκόταν		1	1
ευχαρηστώ		1	1
CBm004			
συνεπια		1	1
φοβούνται	1		1
αντιμετοπίζουμε		1	1
έρχοτε	1		1
CBm005			
διότη		1	1
παρεπιπτόντος	1		1
αναπηχθούν		1	1
καλήτερα		1	1
διμηουργούν		1	1
πρως		1	1
αμηντικά		1	1
σταματίσει	1		1
CBm006			
για (αντί γεια)		1	1
χαρούμενη (οι)	1		1
καταστρέφοντε	1		1
CBm007			
σημμαθήτριες		1	1
παγκόσμια			1
ημέρα			1
λόγο		1	1
ολόκληρως	1		1
η (Η)			1
εξεληγμένων		1	1
CBm008			
κοινωνικό		1	1
κοινωνίες		1	1
Ειρήνη (ει)			1
CBm009			
αλοιώνει		1	1
ταλαιπορημένοι		1	1
ανύμποροι		1	1
αβέβαιοι	1		1
αντιμετοπίσουμε		1	1

CBm010					
πολά		1			1
ελαχιστο		1			1
επίσης		1			1
ποιο		1			1
πολές		1			1
πολά		1			1
επεκτύνουν		1			1
αρχίζουν		1			1
βονβαρδησμοι		1	1		2
πολες		1			1
αλλάξη	1				1
απαγορευτή	1				1
αγαπητή (οί)	1				1
CBm011					
ημέρα				1	1
ειρήνης				1	1
χαρούμενη (οι)	1				1
αναπτυχθοί	1				1
έχοι	1				1
ποιο		1			1
αντιμετώπισης		1			1
συνοψίζοντας		1			1
CBm012	0				
CBm013					
άκρος	1				1
CBm014					
φτώχια		1			1
μίνει		1			1
αναγκαστικά		1			1
CBm015					
Αγαπητοί (α)			1		1
αντιμετοπιστουν		1			1
θανανοση		1			1
συννεπεια		1			1
τυρρανα		2			2
συνεπιων		1			1
διαφήμιση		1			1
ειρηνιστες		1			1
CBm016					
μιλισω	1				1
διευκρινήσω		1			1
ΣΥΝΟΛΟ ΛΑΘΩΝ	20	67	3	9	99

Λέξη / Γραπτό	γραμματικά [G]	ετυμολογικά [E]	φωνολογικά [F]	άλλο [A]	Εμφάνιση
CGm001					
ημέρα				1	1
εριπωμένα		1			1
ξεκινώντας	1				1
σταματήση	2				2
προετιμασίας		1			1
CGm002					
παιδιάς		1			1
συνχησης			1		1
βλάστησης		1			1
συνηδητοποίηση		1			1
επιστημώνων		1			1
δημιουργεία		1			1
ωφέλη		1			1
CGm003					
παγκόσμιας				1	1
ημέρας				1	1
προτίνω		1			1
αντιμετοπισης		1			1
επιρραζουν		2			2
αναπηξη		1			1
κτιρήων		1			1
CGm004					
παγκόσμιας				1	1
ημέρας				1	1
κλήμα		1			1
CGm005					
συνέπιες		1			1
αποζιμιωθούν		1			1
δωθούν		1			1
αποζιμιώσης (εις)	1	1			2
πίνα		1			1
αγχομένο		1			1
αβοήθιτο		1			1
H (η)				1	1
GGm006					
συντρίμα		1			1
μειωση (ει)	1				1
CGm007					
ξενιαστή			1		1
ερύπωση		1			1
προκαλείτε	1				1
πυρινικά		1			1
πλυθησμούς		2			2
προκαταλύψεις		1			1

CGm008			
Ειρήνη (η)			1 1
καμμένα	1		1
CGm009			
αξιότιμη	1		1
Διευθυντή (δ)			1 1
Κυρίες (κ)			1 1
διασφαλήσουμε	1		1
CGm010			
οφείλει	1		1
δυστιχώς	1		1
απολείει	1		1
πιθανών	1		1
συτηση		2	2
πλείτει		2	2
τρόφημα		1	1
δημιουργειας		1	1
ευχαριστισω	1		1
CGm011			
συζητίσουμε	1		1
ασύμαντους		1	1
χυρότερο		1	1
κατασκεβάζονται	1	1	2
κυρίος	1		1
ησχιρές		1	1
πολιτική		1	1
επικρατίσουμε	1		1
λιθή	1	1	2
συμβιβαστικοί (η)	1		1
CGm012			
είνε	1		1
οφεληθεί		1	1
CGm013			
επιλήσουμε		1	1
ευαιστηθοποιουμε			1 1
δημιουργεία		1	1
οδηγείται	1		1
οστόσο		1	1
διαφωρών		1	1
μέσω	1		1
τσακομούς		1	1
παρομειες		1	1
αμβλίνουμε		1	1
ενεργεί (οι)	1		1
CGm014			
Αγαπητοι (α)			1 1
Κυρίες (κ)			1 1

παγκόσμια (Π)				1	1
συμβάλει		1			1
συνήπαρξη		1			1
αναπτύξη	1				1
αλλυλεγγύη		1			1
εκμεταλευση		1			1
CGm015					
ευηκτή		1			1
δυστηχώς		1			1
προοθείται		1			1
CGm016					
διορθωνονται			2		2
μειώνετε	1				1
ΣΥΝΟΛΟ ΛΑΘΩΝ	20	63	3	12	98

Λέξη / Γραπτό	γραμματικά [G]	ετυμολογικά [E]	φωνολογικά [F]	άλλο [A]	Εμφάνιση
AGs001					
γαλαζοπράσινα		1			1
επικίνδυνο		1			1
κενούριους		1			1
μέροι	1				1
μαρέσει			1		1
H (η)				1	1
AGs002					
επίσκεψα (επισκεύτηκα)			1		1
γύρο	1				1
τις (της)	1				1
πάνο	1				1
γραμές		1			1
ήσιχη		1			1
κείματα		1			1
θαλασα		2			2
επιθάνεια			1		1
έπεζα		1			1
πίγαμε		1			1
όλλα		1			1
στεναχορίφηκα	1	1	1		3
AGs003					
πολύ (πολλή)	1				1
δραστηριότητες		1			1
καβούργια			1		1
ψύσουμε		1			1
βρίκαμε		1			1
κρεωπολείο		1			1
της (τις)	1				1
είσουν		1			1

στήλε		1			1
ABs001					
πολύ (πολλή)	1				1
θαλασα		2			2
θαλασούχο		2			2
κολύμπη	1				1
ABs002					
πολή (πολύ)	1				1
σίμερα		1			1
στιν	1				1
κάθο				1	1
εκείνη (εκείνοι)	1				1
βγίκα		1			1
νιάζει		1			1
έφηγα		1			1
σινεχισα		1			1
ABs003					
δίγμα		1			1
περνει		1			1
πολύχρωμες		1			1
εντάζει (αντί πετάζει)				1	1
φανταστικά		1			1
έπησα		1			1
ΣΥΝΟΛΟ ΛΑΘΩΝ	12	32	6	2	52

Λέξη / Γραπτό	γραμματικά [G]	ετυμολογικά [E]	φωνολογικά [F]	άλλο [A]	Εμφάνιση
BGs001					
συμερινη		1			1
το περισσότερο (τον)	1				1
κινιτα		1			1
κληβομένει	1	1	1		3
δοματια		1			1
χωρης	1				1
βγενουν		1			1
κληνονται		1			1
συνεσφίματα		1	1		2
οπος	1				1
βοηθισουν	1				1
δινατοτητα		1			1
απλός(απλώς)	1				1
κουρασμενει	1				1
ομος		1			1
περιστοτερη		1			1
BGs002					
πάντος	1				1
BGs003					

πολύ/πολλοί	1			1
εντάσης (-εις)	1			1
οικογένιες		1		1
στης (αντί στις)	1			1
βρίσημο		1		1
αντιμετοπίσουμε		1		1
ήρεμει (-οι)	1			1
παρέαις	1			1
παρασιρομαστε		1		1
συμβένουν		1		1
η αντί οι	1			1
λόγο (-ω)	1			1
της αντι τις	1			1
νεβρικότιτας	1	1		2
άλλον (-ων)	1			1
τσακόνονται		1		1
σηζιτιφουν		2	1	3
BGs004				
ανταποδόσουν		1		1
γίνετε	1			1
ανχώνουν			1	1
αντιμετοπίζουμε		1		1
BBs001				
οπος		1		1
ολη (-οι)	1			1
η(οι)	1			1
εφιβια		2		2
γινετε	1			1
γωνις	1	1		2
γενεθλιον	1			1
αρεσι	1			1
φοναζουν		1		1
φισικα		1		1
γονις	1			1
στεναχρριουντε	1		1	2
πος	1			1
απαλαγουμε		1		1
αφτο			1	1
βίμα		1		1
γιονγκα			1	1
προσπαθιστε	1			1
μιλισετε	1			1
ειρεμο		1		1
μιλιστε	1			1
BB002				
ειχιο		2		2
διευθυνσει	1			1

ΣΥΝΟΛΟ ΛΑΘΩΝ

32

32

7

0

71

Λέξη / Γραπτό	γραμματικά [G]	ετυμολογικά [E]	φωνολογικά [F]	άλλο [A]	Εμφάνιση
CGs001					
λογο	1				1
αυτή (-οι)	1				1
σκοτινιασει		1			1
κράτοι	1				1
CGs002					
κράτοι	1				1
αποστοδρομοί	1				1
δilhτηριάζεται		1			1
ζώνοντας			1		1
εξελυχθούν		1			1
CGs003					
αφήνοντας	1				1
καταπολεμίσουν	1				1
μεγαλόνωντας	1	1			2
δωθούν		1			1
υοθεσία			1		1
προσπαθίσουν	1				1
CBs001					
είθελα		1			1
συνεποιες		1			1
σκωτονονται		2			2
μειωνετε	1				1
λιγοστένουν			1		1
γίνετε	1				1
περιορησουμε		1			1
προτα		1			1
σκωτονονται		2			2
πολεί (πολλοί)	1				1
μεροι	1				1
CBs002					
πολες		1			1
πολοι		1			1
συνεποιες		1			1
συνεπια		1			1
ανθρωπου		1			1
πολα		1			1
οικογενια		1			1
της (τις)	1				1
κατικιες		1			1
δουλια		1			1
πολη (πολύ)	1				1
αντιμετοπισης		1			1

τσακονεται		1			1
κρατισουν	1				1
μεταξη	1				1
προκαλη	1				1
μηα		1			1
CBs003					
ειρυνης		1			1
συμαντικη		1			1
σταματισουν	1				1
ΣΥΝΟΛΟ ΛΑΘΩΝ	19	27	3	0	49

Έντυπο γονικής συγκατάθεσης

Σκοπός: Επεξεργασία γραπτού λόγου μαθητών/τριών στον ιδιωτικό τομέα (οικοδιδασκαλείου ή φροντιστηρίου) για χρήση υλικού σε μελέτη περίπτωσης στο πλαίσιο διπλωματικής εργασίας.

Προϋπόθεση :
Ότι δεν θίγονται προσωπικά δεδομένα όπως κάθε πληροφορία που αναφέρεται στο παιδί μου, για παράδειγμα το όνομα, τη διεύθυνση της οικίας μας, το τηλέφωνο επικοινωνίας (σταθερό ή κινητό), τα ενδιαφέροντα, επιδόσεις στο σχολείο, κ.ο.κ. σύμφωνα και με την Αρχή Προστασίας Δεδομένων Προσωπικού Χαρακτήρα.

Δηλώνω ότι επιτρέπω να παραχωρηθεί γραπτό υλικό του παιδιού μου για ερευνητικούς σκοπούς στη μεταπτυχιακή φοιτήτρια του Πανεπιστημίου Αιγαίου, Χαραλάμπους Μαρία (ΑΜ:438218030), με τον όρο να τηρηθούν όλες οι άνω προϋποθέσεις.

Αγόρι Κορίτσι

Τάξη Γυμνασίου: Α' Β' Γ'

Καταγωγή: Ελληνική Άλλο

Ονοματεπώνυμο γονέα/κηδεμόνα:

Υπογραφή γονέα/ κηδεμόνα

Ημερομηνία