

# ΠΟΣΟΤΙΚΗ ΑΝΑΛΥΣΗ ΤΟΥ ΛΕΞΙΛΟΓΙΟΥ ΤΗΣ ΤΣΑΚΩΝΙΚΗΣ ΔΙΑΛΕΚΤΟΥ

Ευγενία Τσεμπέλη & Σοφία Στάμου

*Ιόνιο Πανεπιστήμιο*

Every language in the whole world is an abstract concept and more specifically it is considered to be the set of many dialects, both geographical and social. The Modern Greek language is characterized by a long tradition of sociolinguistic differentiation between the natural colloquial language and the archaic ones. These constitute the so-called linguistic varieties of the Greek language and the dialects are the result of these differences. Tsakonian is a language variety which belongs to the geographical dialects. It is characterized by a large deviation in relation to other dialects of our country and in fact in many cases it is considered as a separate language due to the fact that it is incomprehensible to the speakers of Modern Greek. Trying to investigate the structure of languages in general and dialects in particular, but also to interpret them, we turn to Quantitative Linguistics. Using statistical methods, Quantitative Linguistics derives results in order to empirically build on them and express an opinion about a language. In the light of this branch, the Tsakonian dialect is included so that it can be studied as a complementary element in depth. In this way, the goal is to understand some key issues that make this dialect considered difficult and exotic in terms of its pronunciation and its assimilation.

**Keywords:** Modern Greek language, dialect, Tsakonia, Quantitative linguistics

## Εισαγωγή

Με τον όρο διάλεκτο αναφερόμαστε συνήθως στις παραλλαγές μιας γλώσσας η οποία χρησιμοποιείται από ομιλητές μια συγκεκριμένης γεωγραφικής περιοχής. Η έννοια του συγκεκριμένου όρου συνδέεται άμεσα με την έννοια της γλώσσας, καθώς στην καθημερινότητά μας ακούγεται πολλές φορές ότι ένας ομιλητής μιλά τη διάλεκτο της τάδε γλώσσας ή ότι κάποιοι άνθρωποι δεν μιλούν γλώσσα αλλά διάλεκτο.

Σύμφωνα με τον Κοντοσόπουλο (2001), οι διάλεκτοι αποτελούν γλωσσικά συστήματα με κοινή ιστορική καταγωγή που μιλιούνται σε διαφορετικούς τόπους και δε γίνονται κατανοητά από τους ομιλητές της επίσημης γλώσσας.

Η ελληνική γλώσσα αν και παρέμεινε ενιαία από την αρχαιότητα μέχρι και σήμερα, ποτέ δεν μιλήθηκε ομοίμορφα σε ολόκληρο τον ομόγλωσσο γεωγραφικό χώρο

με αποτέλεσμα να διαμορφωθούν έτσι τοπικές διαλεκτικές παραλλαγές. Πιο συγκεκριμένα διαφοροποιημένη είναι η αρχαία ελληνική που φτάνει μέχρι και τους κλασσικούς αιώνες με διάφορες παραλλαγές να επικρατούν και να ομιλούνται στις πόλεις-κράτη. Οι διάλεκτοι που επικράτησαν τότε ήταν η ιωνική, η δωρική, η αττική και η αιολική. Το ίδιο διαφοροποιημένη είναι και η σύγχρονη ελληνική. Ως διαλέκτους σήμερα θεωρούμε την Ποντιακή, την Καππαδοκική, την Τσακωνική και την Κατωιταλική, Κρητική και Κυπριακή.

Στη σύγχρονη γλωσσολογία οι διάλεκτοι εντάσσονται στη γενικότερη μελέτη των γλωσσικών ποικιλιών μιας έννοιας που θεωρείται ότι αποδίδει καλύτερα την ανομοιογένεια της γλωσσικής πραγματικότητας απ' ό,τι η παλαιότερη αντίληψη που ήθελε τη γλώσσα ένα ομοιογενές σύστημα σημείων. Το φάσμα που μπορεί να καλύψει η γλωσσολογία είναι πολύ μεγάλο και όσον αφορά τη μελέτη των διαλέκτων εφαρμόζεται ένας υπό-κλάδος της κοινωνιογλωσσολογίας που ονομάζεται διαλεκτολογία.

Μια από τις πολλές διαλέκτους της ελληνικής γλώσσας είναι και η τσακωνική διάλεκτος, η οποία μάλιστα κατάγεται από την αρχαία δωρική διάλεκτο της Λακωνίας και εντάσσεται στις γλώσσες υπό εξαφάνιση σύμφωνα με την Unesco<sup>1</sup>. Η τσακωνική διάλεκτος εμφανίζει σημαντικές διαφορές από την κοινή νεοελληνική γλώσσα και τις υπόλοιπες διαλέκτους, με αποτέλεσμα να ηχεί για τους μη φυσικούς ομιλητές ως μια ακατανόητη (σχεδόν εξωτική) γλώσσα. Η τσακωνική διάλεκτος ομιλήθηκε και μιλιέται μέχρι και τις μέρες μας στην περιοχή της Τσακωνιάς, η οποία γεωγραφικά καταλαμβάνει το μεγαλύτερο μέρος της περιοχής της Κυνουρίας, που βρίσκεται νοτιοανατολικά του νομού Αρκαδίας. Η μορφολογία του εδάφους που καθιστούσε μέχρι το πρόσφατο παρελθόν δύσβατη και άρα απομονωμένη την περιοχή συνέβαλε καθοριστικά στη διατήρηση της τοπικής αυτής γλώσσας.

Για τη διάλεκτο έχουν γίνει αρκετές έρευνες από Έλληνες και ξένους επιστήμονες και γλωσσολόγους. Ένας από τους πρώτους είναι ο Θ. Οικονόμου που εκδίδει το την πρώτη γραμματική της τσακωνικής διαλέκτου (1870), ενώ ο Μ. Δέφνερ εκδίδει την Τσακωνική Γραμματική (Deffner 1881) και το Λεξικόν της Τσακωνικής διαλέκτου (Δέφνερ 1923). Επιπλέον, ο Γ. Χατζιδάκης εκδίδει μελέτες και άρθρα με αναφορά στην Τσακωνική διάλεκτο (Hatzidakis 1892· Χατζιδάκης 1905) ενώ ο Η. Pernot, δημοσίευσε το βιβλίο «Εισαγωγή στη μελέτη Τσακωνικής Διαλέκτου». Ο Κωστάκης εκδίδει το τρίτομο λεξικό της Τσακωνικής διαλέκτου (1986α· 1986β· 1987) σε συνεργασία με την Ακαδημία Αθηνών, ενώ το 1954 ιδρύεται το Αρχείο Τσακωνιάς<sup>2</sup>, το οποίο έχει εκδόσει πλήθος βιβλίων που αναφέρονται στη γλώσσα καθώς και στα ήθη και τα έθιμα της Τσακωνιάς.

<sup>1</sup> <http://www.unesco.org/languages-atlas/> (ημερομηνία πρόσβασης: 10.12.2021).

<sup>2</sup> <http://www.tsakonianarchives.gr/> (ημερομηνία πρόσβασης: 11.12.2021).

Σημαντικό έργο του Αρχείου Τσακωνιάς είναι μεταξύ άλλων και η ψηφιοποίηση έντυπου υλικού αλλά και η οργάνωση δράσεων για την καταγραφή και διάσωση της ιδιαίτερης αυτής διαλέκτου.

Στο παρόν άρθρο θα παρουσιάσουμε μια πρότυπη μελέτη του λεξιλογίου της τσακωνικής διαλέκτου υπό το πρίσμα της ποσοτικής γλωσσολογίας, δηλαδή του κλάδου της γενικής γλωσσολογίας που χρησιμοποιεί στατιστικές μεθόδους για τη μελέτη της γλώσσας. Στόχος μας είναι να καταδείξουμε ποσοτικά την κατανομή της συχνότητας των γραμματικών κατηγοριών, του μέσου μήκους των λέξεων, της συχνότητας των γραμμάτων και των λέξεων της τσακωνικής διαλέκτου, ώστε να συμβάλλουμε αφενός στις προσπάθειες που συντελούνται για την πλήρη περιγραφή της τσακωνικής διαλέκτου και αφετέρου να δώσουμε το έναυσμα για τη συγκριτική αξιολόγηση του λεξιλογίου της τσακωνικής διαλέκτου με το λεξιλόγιο είτε άλλων διαλέκτων είτε της κοινής ελληνικής γλώσσας. Για την εκπόνηση της μελέτης μας χρησιμοποιήσαμε το λεξιλόγιο της τσακωνικής διαλέκτου όπως έχει καταγραφεί στο λεξικό του Deffner και εφαρμόσαμε τις ποσοτικές μεθόδους γλωσσολογικής έρευνας που έχουν υιοθετηθεί στην έρευνα των Μικρός κ. ά. (2003).

## 2. Οι Τσάκωνες και η Τσακωνιά

### 2.1. Ιστορία της Τσακωνιάς

Τα πρώτα δείγματα κατοίκησης και ανθρώπινης παρουσίας σύμφωνα με τους αρχαιολόγους στην ευρύτερη περιοχή της Τσακωνιάς ξεκινάει ήδη από την Μυκηναϊκή εποχή. Σύμφωνα με πολλές ιστορικές μαρτυρίες οι Τσάκωνες θεωρούνται άμεσοι απόγονοι των Δωριέων οι οποίοι ζούσαν απομονωμένοι. Γι' αυτό και κατάφεραν να διατηρήσουν αναλλοίωτα τα ήθη τους, τα έθιμα τους και τη διάλεκτό τους.

Η Κυνουρία και ιδίως όσες περιοχές βρίσκονται κοντά στη θάλασσα αποτέλεσαν «το μήλον της Έριδος» ανάμεσα στους Λακεδαιμόνιους, τους Αργείους και τους Αθηναίους. Το 195 π. Χ. ύστερα από επέμβαση του Υπάτου Ρωμαίου Τίτου Κόιντου Φλαμίνιου, η περιοχή δίδεται στους Αργείους. Ωστόσο, οι Σπαρτιάτες επιμένουν και επιτυγχάνουν να κρατήσουν τις Πρασιές, οι οποίες εντάσσονται στο «κοινόν των Λακεδαιμονίων» μαζί με άλλες είκοσι τρεις λακωνικές πόλεις.

Την περίοδο της Ρωμαϊοκρατίας οι Τσάκωνες συνεχίζουν να μένουν απομονωμένοι. Οι βασικές τους ασχολίες είναι οι αγροτικές δουλειές ή καταφεύγουν στα βουνά

και ασχολούνται με την κτηνοτροφία. Εκείνη την περίοδο οι φοβερές επιδρομές των Σλάβων στην Ελλάδα και την Πελοπόννησο επηρεάζουν τους Τσάκωνες. Αυτό τους αναγκάζει να μεταναστεύσουν άλλοι στα ανατολικά και στις πλαγιές του Πάρωνα και άλλοι στον Πόντο, στα νησιά του Αιγαίου και στην Κάτω Ιταλία. Οι Σλάβοι καθυπόταξαν ολόκληρη την Πελοπόννησο, την Τσακωνιά όμως δεν κατάφεραν να την υποδουλώσουν. Για τους Τσάκωνες ήδη από την εποχή του Ιουστινιανού (527 μ. Χ) έχει ειπωθεί σύμφωνα με γραπτές αναφορές βυζαντινών συγγραφέων ότι είναι γνήσιοι απόγονοι των Δωριέων και περιγράφονται ως «επίλεκτα σώματα των καστροφυλάκων» και της σωματοφυλακής των βυζαντινών αυτοκρατόρων. Γι' αυτό και στο Βυζάντιο υπάρχει το αξίωμα του «Στρατοπεδάρχη των Τσακώνων». Επιπλέον, πολλοί Τσάκωνες αξιοποιήθηκαν στον Βυζαντινό στόλο.

Κατά τη Φραγκοκρατία η Τσακωνιά παραμένει πιστή στο Βυζάντιο και δίνει σκληρούς αγώνες εναντίων των κατακτητών. Σε πολλές ιστορικές πηγές αναφέρεται η υποδούλωση του Μοριά στους Φράγκους το 1205 μ.Χ., ο οποίος στη συνέχεια διαχωρίστηκε σε φέουδα και το κομμάτι της Τσακωνιάς κυριαρχείται από τον Γκυ ντε Νιβελαί. Ο Γοδεφρείδος και ο Γουλιέλμος Βιλλεαρδουίνος τιμωρούν αμείλικτα τους ανυπότακτους Τσάκωνες και χτίζουν τα κάστρα των Γεροθρών και Μιστρά για να τους ελέγχουν. Αυτή την περίοδο, γύρω στο 1293 μ. Χ., γίνεται η πρώτη αναφορά του Λεωνιδίου σε χρυσόβουλο του Ανδρόνικου Β' Παλαιολόγου. Εκεί, για πρώτη φορά, αναφέρεται ο ναός του Λεωνίδα, από τον οποίο πήρε το όνομά της η περιοχή.

Μετά από πολλές συγκρούσεις η Τσακωνιά περνά στα χέρια των Τούρκων. Κατά την Πρώτη Τουρκοκρατία (1453–1715 μ. Χ.) οι Τσάκωνες βοηθούν τους Ενετούς στον αγώνα τους κατά του Τουρκικού ζυγού, ενώ ξεκινούν εμπόριο βουτύρου στην Ευρώπη και εκμεταλλεύονται αλατωρυχεία στα σύνορα Αυστρίας και Πρωσίας. Στη συνέχεια, η περιοχή περνά για λίγο στα χέρια των Βενετσιάνων, όμως μετά το 1715 μ.Χ. κατακτάται ξανά από τους Τούρκους και η Τσακωνιά σταδιακά γνωρίζει οικονομική και κοινωνική ανάπτυξη. Την περίοδο αυτή οι Τσάκωνες αρχίζουν να χτίζουν αρχοντικά και πυργόσπιτα στις ιδιαίτερες πατρίδες τους και προσφέρουν χρήματα για την Ελληνική Επανάσταση του 1821 μ. Χ. Σημαντικές φυσιογνωμίες είναι ο Κωνσταντίνος Τσικαλιώτης και ο Κώστας Χατζήπαναγιώτης. Ανάμεσα στις περιοχές της Τσακωνιάς, ο Πραστός αποκτά πλούτο και μετατρέπεται σε ισχυρό οικονομικό και διοικητικό κέντρο, ενώ γίνεται η πρώτη πρωτεύουσα της Τσακωνιάς. Το 1826 μ. Χ. ο Ιμπραήμ μαζί με τους Τουρκοαιγυπτίους εισβάλλουν στον Πραστό και τον καταστρέφουν ολοκληρωτικά. Οι ξεριζωμένοι Τσάκωνες αναγκάζονται να εγκαταλείψουν τον Πραστό και να εγκατασταθούν στο Λεωνίδιο, το οποίο ο Κολοκοτρώνης ονόμαζε «Σίγουρο Τόπο». Έτσι το 1845 μ. Χ. το Λεωνίδιο το οποίο ποτέ δεν υπέστη τις βιαιότητες του Ιμπραήμ λόγω της γεωγραφικής του θέσης ορίζεται η νέα πρωτεύουσα της Τσακωνιάς.

Την περίοδο του μεσοπολέμου παρατηρούνται έντονα μεταναστευτικά ρεύματα προς την Κωνσταντινούπολη αλλά και προς την Κυνουρία με αποτέλεσμα να αυξηθεί ο πληθυσμός του Λεωνιδίου. Μετά την πολεμική δεκαετία του 1940 μ. Χ. πολλοί κάτοικοι στρέφονται προς την Αθήνα και την υπερπόντια μετανάστευση. Από το 1970 μ. Χ. και μετά παρατηρείται αύξηση του πληθυσμού καθώς επιστρέφουν πολλοί από αυτούς που είχαν μεταναστεύσει στην Αυστραλία. Οι κάτοικοι αρχίζουν να ασχολούνται με την οικοδομή και την καλλιέργεια περιβολιών. Από τότε μέχρι και σήμερα αρκετοί Λεωνιδιώτες ασχολούνται με τη γεωργία.

## 2.2. Η Τσακωνική διάλεκτος

Η Τσακωνική διάλεκτος είναι ελληνογενής διαλεκτική ομάδα που μιλιέται στην περιοχή της νότιας Κυνουρίας της Αρκαδίας στην Πελοπόννησο. Ανήκει στην υπολειμματική δωρική ζώνη της Νέας Ελληνικής, δηλαδή έχει προέλθει από τη Δωρική διάλεκτο της Αρχαίας ελληνικής και γι' αυτό το λόγο χαρακτηρίζεται ως αρχαιοπρεπής. Οι πρώτες επίσημες αναφορές για την ύπαρξη μιας ιδιαίτερης τσακωνικής διαλέκτου γίνεται από τον συγγραφέα Μάζαρις τον 15<sup>ο</sup> αιώνα ο οποίος δεν καταφέρνει να καταγράψει τσακωνικές λέξεις. Ουσιαστικά η πρώτη μνεία για την διάλεκτο γίνεται το 1668 από τον Εβλιά Τσελεμπί, ο οποίος κατάφερε να καταγράψει μερικές λέξεις, συνολικά 35. Παρά το μικρό αριθμό των λέξεων όμως αντικατοπτρίζεται η τσακωνική μέχρι ένα βαθμό. Μάλιστα ο Τσελεμπί τονίζει ότι οι Τσάκωνες μιλούν μια περίεργη γλώσσα η οποία δεν είναι ούτε ιταλικά ούτε ελληνικά και χρειάζεσαι διερμηνέα για να συνεννοηθείς μαζί τους. Ο Γερμανός φιλόλογος και γλωσσολόγος Μιχαήλ Δέφνερ είναι ο πρώτος που συντάσσει επίσημα ένα Λεξικό της Τσακωνικής διαλέκτου και δημοσιεύτηκε το 1923 και ακολουθεί αρκετά χρόνια αργότερα το Λεξικό του Θανάση Κωστάκη (1986α· 1986β· 1987).

Η Τσακωνική διάλεκτος έτσι όπως ομιλείται και έχει επικρατήσει με την πάροδο του χρόνου χωρίζεται σε δυο κατηγορίες, στη νότια και τη βόρεια. Η νότια ομιλείται στα χωριά Λεωνίδιο, Πραγματευτή, Μέλανα, Σαπουνακείκα και Τυρός, η βόρεια ομιλείται στα χωριά Πραστός, Καστάνιτσα, Σίταινα, Άγιος Ανδρέας. Σύμφωνα με τον Κωστάκη (1951) όμως, μπορούμε να διακρίνουμε και μια τρίτη κατηγορία αυτής της διαλέκτου και αυτή είναι τα τσακωνικά της Προποντίδας. Αυτά χρησιμοποιούνται σε κάποια χωριά στα μικρασιατικά παράλια όπου με μετά την μικρασιατική καταστροφή έρχονται στην Ελλάδα. Αυτή η μορφή της τσακωνικής όμως είναι επηρεασμένη από τα Θρακικά τα οποία θυμίζουν περισσότερα το Νέα Ελληνικά.

Οι εντυπωσιακές διαφορές που παρουσιάζει η τσακωνική διάλεκτος από τη Νέα Ελληνική αλλά και από άλλες διαλέκτους την καθιστά για πολλούς ομιλητές ως μια «εξωτική» γλώσσα. (Τζιτζιλής 2000). Η γεωγραφική θέση της Τσακωνιάς επίσης είναι ένας ακόμα λόγος που καθιστούν αυτή τη διάλεκτο πιο δύσκολη, καθώς η μορφολογία του εδάφους είναι αρκετά κακοτράχαλη, επομένως η κοινωνία ήταν κλειστή και απομονωμένη. Η ανάπτυξη παρουσιάζεται στα μέσα του προηγούμενου αιώνα όπου δημιουργείται οδικό δίκτυο και παρουσιάζεται οικονομική και όχι μόνο ανάπτυξη στην περιοχή.

Στις μέρες μας, η διάλεκτος όπως αναφέρει ο Μηνάς (2004), «ψυχορραγεί» καθώς η ανάπτυξη της περιοχής και η επαφή με ξένους καθιστά δύσκολη την επικοινωνία. Ειδικότερα οι νέοι ότι δεν τους προσφέρει τίποτα και ιδιαίτερα σε ότι έχει σχέση με την επαγγελματική αποκατάσταση. Παρόλα αυτά γίνονται κατά καιρούς προσπάθειες αναβίωσης μέσα από ειδικά προγράμματα εκπαίδευσης και ενημέρωσης. Πλέον έχει χαρακτηριστεί από την UNESCO ως ένα είδος προς εξαφάνιση.

Τα κυριότερα χαρακτηριστικά της τσακωνικής διαλέκτου είναι:

1. *a /a/* από *a /a:/* της δωρικής και ινδοευρωπαϊκής αντί για το *η /i/* της νεοελληνικής (NE), το οποίο είναι εξέλιξη του *η /e:/* της αττικής-ιωνικής (από προγενέστερο /a:/) π. χ. *α μάτη* [a máti] 'ή μητέρα' από δωρική *ά μάτηρ* [ha: má:te:r], σε αντίθεση με *η μητέρα* [i mitéra], το οποίο παράγεται από το *ή μήτηρ* [he: mé:te:r] της αττικής-ιωνικής κοινής.
2. *σ /s/* σε μερικές λέξεις αντιστοιχεί στο *θ /tʰ/* της αρχαίας ελληνικής (ΑΕ) αντί για *θ /θ/* στη NE — από το (πιθανό) */s/* της λακωνικής διαλέκτου, π. χ. *ένι ασού* [éni asú] 'σουρώνω' πρβ. ΑΕ *ήθῶ* (-έω), (πιθανό) δωρ. *άθῶ*, λακων. *άσῶ*. Συγκρίνετε τις λακωνικές επιγραφές με ANEΣΗΚΕ [anes:ke] 'αφοσίωσε' αντί για το συνηθέστερο ANEΘΗΚΕ [anetʰe:ke] ή αργότερα [aneθe:ke]. Άλλα παραδείγματα: *ένι κασήμενε* [éni kasímene] 'κάθομαι' πρβ. ΑΕ *καθήμενος*, *σάτη* [sáti] 'κόρη' πρβ. ΑΕ *θυγάτηρ*.
3. *ρ /r/* σε τελική θέση αντίστοιχο με το τελικό σίγμα */s/* της NE και των περισσότερων διαλέκτων της ΑΕ (φαινόμενο γνωστό ως **ρωτακισμός**). Στα τσακωνικά το τελικό *ρ /r/* απαντάται μόνο εφόσον αρχίζει η επόμενη λέξη με φωνήεν, αλλά το φαινόμενο μάλλον κατάγεται από τη μεταγενέστερη λακωνική. Πρβ. το τσακωνικό *τῆϊρ αμέρε* [tʰi:r amére] 'τρεις ημέρες' με το NE *τρεις ημέρες* [tris iméres] από τη μία πλευρά, και λακωνικές επιγραφές της ύστερης ρωμαϊκής περιόδου στις οποίες βρίσκουμε NEIKAAP [ni:ka:a:r] 'έχοντας νικήσει' αντί για το NEIKHΣΑΣ [ni:ke:sas] από την άλλη.
4. Κλίση ρημάτων — ο ενεστώτας και ο παρατατικός της οριστικής σχηματίζονται περιφραστικά με το ρήμα 'είναι' + μετοχή ενεστώτα (ενεργητική ή παθητική), η οποία

κλίνεται ως προς το γένος και τον αριθμό. Η μετοχή αυτή κατάγεται από τη μετοχή ενεστώτα της ΑΕ η οποία κλινόταν και είχε και παθητική φωνή, σε αντίθεση με εκείνη της ΝΕ που έχει μόνο ενεργητική μορφή και είναι άκλιτη. Επιπλέον, αυτές οι σύνθετες μορφές έχουν υποκαταστήσει εντελώς τον ενεστώτα και τον παρατατικό οριστικής της ΑΕ (εκτός από το 'είμαι'), αν και υπάρχει ακόμα η υποτακτική του ενεστώτα στα τσακωνικά.

5. ντ [nd] < [d:] δδ και μπ [mb] < [b:] ββ της λακωνικής (και άλλων διαλέκτων της ΑΕ, συχνά αποτέλεσμα αποκοπής): *έρι σερίντου* [éri serindu] 'θερίζω', *έρι λάμπαινου* [éri k<sup>h</sup>ambénu] 'κατεβαίνω'. Πρβ. *Θερίδδεν* (Αριστοφάνη «Αχαρνείς» στ. 947 — ομολογουμένως στη βοιωτική διάλεκτο), *κάβασι κατάβηδι*. *Λάκωνες* (Ησύχιος), *μικχιδδομένων* (επιγραφή της Σπάρτης (αφιέρωμα στην Άρτεμη Ορθία, Β' αιών. μ. Χ.), εναλλακτική μορφή του *μικχιζίζομαι*).
6. Μεμονωμένες λέξεις με β [v] από ΑΕ δίγαμμα F [w]: *βάννε* [váne] 'αρνί' < ΑΕ \**Fάρνος*, πρβ. *Φαρήν* (επιγραφή Γόρτυνος Κρήτης), *ναβελέ* [davelé] 'δαυλί' < ΑΕ \**δαFελός*, πρβ. Ησύχιος *δαβελός*, *δαλός*. *Λάκωνες*.
7. Πτώση του σ μεταξύ δύο φωνηέντων σε συγκεκριμένες ρηματικές καταλήξεις: *ορούα* [orúa] 'βλέποντας' (θηλ. μετοχή) < Λακωνική *όρωα* (πρβ. *Νικάβας* σε λακωνικές επιγραφές), *αλήωι* [alíoi] '[να] πουν' < ΑΕ *λαλήσωσι*, *ερέκαϊ* [erékai] 'βρήκανε' < ΑΕ *εύρηκασι*.
8. Δασέα σύμφωνα (άηχα κλειστά) από συμπλέγματα διάφορων συμφώνων (π. χ. συριστικό + κλειστό, έρρινο + δασύ κλειστό) της ΑΕ.
9. Προσωπικές αντωνυμίες πολύ διαφορετικές από τις αντίστοιχες αντωνυμίες της νεοελληνικής:  
*εζού* [ezú] 'εγώ', *ενίου* [eríu] 'εμένα' (πρβ. ΑΕ *έμεϊο*, δωρ. *έμευ*, *έμιάω*)  
*εκιού* [ecú] 'εσύ' (ΑΕ δωρ. *τυ*)  
*νι* [ni] 'τον/τη/το' — αιτιατική γ' προσώπου εν., π. χ. *νι οράκα* [ni oráka] 'τον/την/το είδα' (ΑΕ δωρ. *νιν*)  
*σι* [si] 'του/της/του' — γενική γ' προσώπου εν., π. χ. *το καμπζί σι* [to kabzí si] 'το παιδί του/της' (ΑΕ *σφι*)  
*νάμου* [námu] 'μας' — γενική και αιτιατική/δοτική α' προσώπου πληθ., π. χ. *έν' ορού νάμου* [én orú námu] 'μας βλέπει, α γρούσσα νάμου' [a grússa námu] 'η γλώσσα μας' (ΑΕ δωρική *άμων* > *αμούν* > *νάμου*)



εμού [emú] ‘εσείς’, π. χ. *εμού και ‘ε’ έγγουντε*; [emú ca t<sup>h</sup> éhgunde] ‘εσείς πού πάτε;’ (ΑΕ δωρ. *ύμέζ > ύμέρ > ύμέ > ουμέ > (μετάθεση) εμού*)

σι [si] ‘τους/τις/τα’ — αιτιατική/δοτική γ’ προσώπου πληθ., π. χ. *ένι νίου σι, μα όνι ορού σι* [éni níu si ma óri oú si] ‘τους/τις/τα ακούω μα δεν τους/τις/τα βλέπω’ (ΑΕ σφε)

σου [su] ‘τους’ — γενική γ’ προσώπου πληθ., π. χ. *ετήνα έν’α τσέα σου* [etína én a dzéa su] ‘εκείνο είναι το σπίτι τους’ (ΑΕ σφών).

10. Συγκρητισμός των πτώσεων — πολλά ουσιαστικά και επίθετα έχουν μόνο μια πτώση ανά αριθμό, π. χ. *ο/τον/του άντε* [áde] ‘ψωμί, *οι/τουρ άντου* [ádu] ‘ψωμιά’ (ΑΕ *άρτος*). Η τάση αυτή απαντά και στη νεοελληνική, αλλά είναι σε πιο προχωρημένο στάδιο στα τσακωνικά.
11. ε /e/ αντί για το ο /o/ της νέας και αρχαίας ελληνικής, ύστερα από ουρανικό ή οδο-ντικό σύμφωνο (ή και το φωνήεν ι) στην κλίση του ενικού αρσενικών και ουδέτερων ονομάτων, επιθέτων, μετοχών και αντωνυμιών. Π. χ. *άλλε* [ále] ‘άλλος’, *ουρανέ* [urané] ‘ουρανό’, *ορατέ* [oraté] ‘ιδωμένος’ (< ΑΕ *όρατός*), *ατσέ* [atse] ‘μεγάλος’ (< ΑΕ *άδρός*). Πρβ. *άτσωπο* [átsofo] ‘άντρας’ και *λιούκο* [líko] ‘λύκος’ έπειτα από χειλικό και υπερωικό σύμφωνο αντίστοιχα. Πιθανότατα η αλλαγή έγινε με την επίδραση του «μπροστινού» (φωνητικά) στοιχείου που προηγήθηκε και του ρ /r/ που ακολούθησε.
12. Τροπή του ρ [r] σε ζ [z] μπροστά από ι [i]. Π. χ. *μακάζι* [makázi] ‘μακάρι, ζάτσι [záti] ‘ρύακι, *έζιφο* [ézifo] ‘κατσίκι’ (< ΑΕ *έριφος*). Στο βόρειο ιδίωμα η προφορά είναι [rɜ] και εικάζεται ότι αυτή ήταν η γενικότερη προφορά και στα δυο ιδιώματα στο 19ο αιώνα. Οι άντρες κατά κανόνα δεν χρησιμοποιούν το ζ [z], προτιμώντας το ελληνικό ρ.

### 3. Ποσοτική γλωσσολογία

Η ποσοτική γλωσσολογία (QL) είναι μια υποκατηγορία της γενικής γλωσσολογίας και, πιο συγκεκριμένα, της μαθηματικής γλωσσολογίας. Η ποσοτική γλωσσολογία ασχολείται με την εκμάθηση γλωσσών, την αλλαγή γλώσσας και την εφαρμογή καθώς και τη δομή των φυσικών γλωσσών. Επίσης, διερευνά γλώσσες χρησιμοποιώντας στατιστικές μεθόδους. Ο πιο απαιτητικός στόχος του είναι η διατύπωση των γλωσσικών νόμων και, τελικά,



μιας γενικής θεωρίας της γλώσσας υπό την έννοια ενός συνόλου αλληλένδετων γλωσσικών νόμων.

Η ποσοτική γλωσσολογία βασίζεται εμπειρικά στα αποτελέσματα των στατιστικών για τις γλώσσες, ενός πεδίου που μπορεί να ερμηνευτεί ως στατιστικό στοιχείο μια γλώσσας ή στατιστικό στοιχείο οποιουδήποτε αντικειμένου. Ο τομέας αυτός δεν συνδέεται απαραίτητα με σημαντικές θεωρητικές φιλοδοξίες. Η ορθολογική και η υπολογιστική γλωσσολογία είναι άλλοι τομείς που παρέχουν σημαντικά εμπειρικά στοιχεία.

Για να μπορέσουν να εφαρμοστούν οι στατιστικές έχουν αναπτυχθεί κάποιοι νόμοι πάνω στους οποίους βασίζονται. Με τον όρο «νόμος» στην ποσοτική γλωσσολογία εννοούμε κάποιες θεωρητικές παραδοχές που μπορούν να διατυπωθούν μαθηματικά.

Οι γλωσσικοί νόμοι είναι στατιστικές κανονικότητες που εμφανίζονται σε διάφορες γλωσσικές κλίμακες (π. χ. φωνήματα, συλλαβές, λέξεις ή προτάσεις) που μπορούν να διατυπωθούν μαθηματικά και έχουν συναχθεί από ορισμένες θεωρητικές παραδοχές. Απαιτείται επίσης να έχουν δοκιμαστεί επιτυχώς μέσω της χρήσης δεδομένων, δηλαδή να μην έχουν απορριφθεί από εμπειρικά στοιχεία. Παρακάτω θα αναφέρουμε συνοπτικά τους νόμους με τους οποίους ασχολείται η ποσοτική γλωσσολογία:

1. Νόμος του Zipf: η συχνότητα των λέξεων είναι αντιστρόφως ανάλογη της κατάταξής τους στους καταλόγους συχνότητων.
2. Νόμος του Hear: περιγράφει τον αριθμό των διακριτών λέξεων σε ένα έγγραφο ή ένα σύνολο εγγράφων ως συνάρτηση του μήκους του εγγράφου.
3. Νόμος της σύντμησης: βασίζεται στο νόμο του Zipf και δηλώνει ποιοτικά ότι όσο πιο συχνά χρησιμοποιείται μια λέξη τόσο πιο μικρή είναι.
4. Νόμος των Menzerath και Altman: ο νόμος αυτός αναφέρει ότι τα μεγέθη των συστατικών ενός έργου γενικά και ειδικότερα ενός κειμένου μειώνονται όσο αυξάνεται το έργο. Πιο συγκεκριμένα εξέτασαν ότι όσο πιο μεγάλη είναι η διάρκεια μιας λέξης σε συλλαβές ή χαρακτηριστές τόσο συντομότερες είναι οι συλλαβές ή οι λέξεις σε ήχους.
5. Νόμος της διαφοροποίησης: οι γλωσσικές κατηγορίες όπως είναι τα μέρη του λόγου εμφανίζονται με διάφορες μορφές και μπορεί να αποδειχθεί ότι οι συχνότητες εμφάνισης τους στα κείμενα μπορούν να ελεγχθούν από τους νόμους.
6. Νόμος του Martin: ο νόμος αυτός αφορά τις λεκτικές αλυσίδες που προκύπτουν από την αναζήτηση του ορισμού σε ένα λεξικό. Όλοι οι ορισμοί σχηματίζουν

μια ιεραρχία όλο και πιο γενικών όρων. Μεταξύ αυτών των επιπέδων ιεραρχίας υπάρχει ένας αριθμός νόμιμων σχέσεων.

7. Νόμος συνόλου κειμένου: οι γλωσσικές μονάδες (π. χ. φωνήματα, συλλαβές, λέξεις, προτάσεις) εμφανίζουν μια συγκεκριμένη κατανομή συχνότητας σε εξίσου μεγάλα τμήματα κειμένου.

### 3.1. Νόμος του Zipf

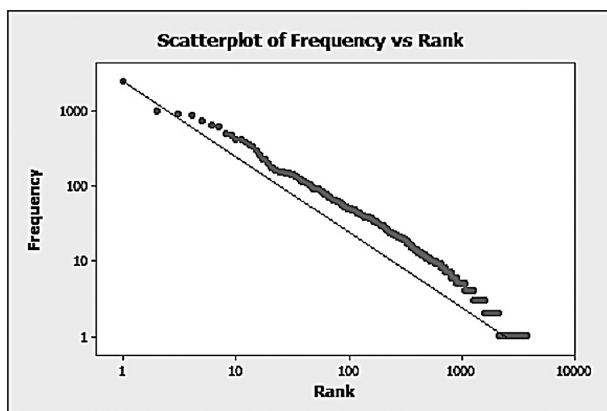
Ο νόμος του Zipf πήρε το όνομά του από τον Αμερικανό γλωσσολόγο George Kingsley Zipf (1902–1950) και είναι ένας εμπειρικός νόμος που αναφέρεται στο γεγονός ότι πολλοί τύποι των δεδομένων που μελετήθηκαν και στις φυσικές και στις κοινωνικές επιστήμες, ακολουθούν μια κατανομή γνωστή ως νόμος των δυνάμεων (power law).

Σύμφωνα με το Oxford Dictionary (2015) νόμος των δυνάμεων (powerlaw) είναι «μια σχέση μεταξύ δυο ποσοτήτων έτσι ώστε η μια να είναι ανάλογη με μια σταθερή δύναμη της άλλης». Ο ίδιος ο Zipf ασχολήθηκε με το να προβλέψει τη συχνότητα της εμφάνισης των λέξεων μέσα σε ένα κείμενο. Δηλώνει δε ότι «αν οι λέξεις ταξινομηθούν κατά φθίνουσα σειρά του αριθμού εμφάνισής τους σε ένα σχετικά μεγάλο κείμενο, τότε η θέση/σειρά μιας λέξης σε αυτόν τον κατάλογο όταν πολλαπλασιάζεται με τη συχνότητα εμφάνισής της είναι ίση με μια σταθερά». Η εξίσωση για αυτή τη σχέση είναι:

$$r \times f = k$$

όπου  $r$  είναι η θέση/σειρά της λέξης,  $f$  είναι η συχνότητα και  $k$  είναι η σταθερά. Για παράδειγμα αν μέσα σε ένα αρκετά μεγάλο κείμενο μια λέξη που βρίσκεται στην θέση 1 (π. χ. η λέξη the) εμφανίζεται 1000 φορές, τότε η λέξη της θέσης 2 (π. χ. η λέξη of) θα εμφανίζεται περίπου 500 φορές, η λέξη της θέσης 3 (π. χ. η λέξη and) θα εμφανίζεται περίπου 300 φορές.

Η εγκυρότητά του έχει παρατηρηθεί σε ένα μεγάλο εύρος φαινομένων, συμπεριλαμβανομένων των φυσικών γλωσσών, οικονομικών, οικολογικών συστημάτων και στατιστικών πρόσβασης σε δικτυακούς τόπους. Στην ελληνική γλώσσα έχει εφαρμοστεί ο πρώτος νόμος του Zipf για να αξιοποιηθεί η πρώτη έκδοση του Εθνικού Θησαυρού της Ελληνικής Γλώσσας που αποτελείται από 13 εκατομμύρια λέξεις και της δεύτερης που είναι εμπλουτισμένη και αποτελείται από 33 εκατομμύρια λέξεις.



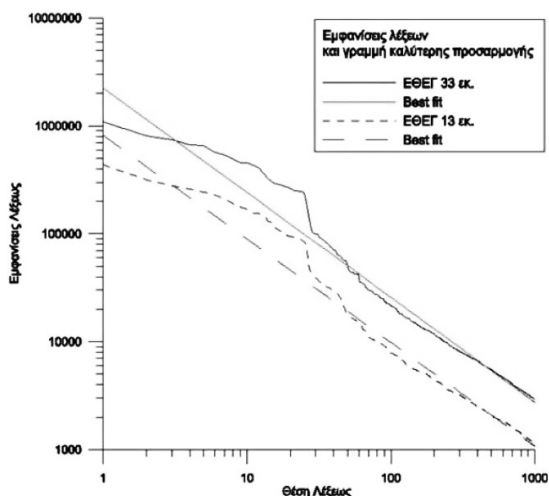
**Εικόνα 1:** Διάγραμμα διασποράς για το νόμο Zipf

Η συχνότητα των λέξεων για τις 1000 πιο συχνές λέξεις ακολουθεί αρκετά πιστά τον νόμο του Zipf, τόσο στην έκδοση 13 εκ. λέξεων του ΕΘΕΓ, όσο και σε αυτήν των 33 εκ. λέξεων.

Παρακάτω παρουσιάζονται οι καμπύλες της συχνότητας των λέξεων. Για τον ΕΘΕΓ των 33 εκ. λέξεων, η παράμετρος  $a$  της εξίσωσης  $1 (\log(p) B a \log(i) i b p^i = a \Rightarrow i = -$ , με  $a \approx 1$ ) παίρνει την τιμή 0,96 ενώ για τον ΕΘΕΓ των 33 εκ. λέξεων την τιμή 0,97. Και οι δύο αυτές τιμές δεν είναι ακριβείς διότι τα πρώτα 20 σημεία του διαγράμματος αποκλίνουν αρκετά από την ευθεία, όμως το σημαντικό στοιχείο των αριθμών αυτών είναι ότι βρίσκονται πολύ κοντά στην μονάδα, όπως ακριβώς θα περιμέναμε. Επίσης, τόσο η μορφή των καμπύλων, όσο και παράμετροι των ευθειών καλύτερης προσαρμογής είναι παρόμοιοι για τις δύο μορφές του ΕΘΕΓ, έστω και αν η συνολική ποσότητα των λέξεων έχει σχεδόν τριπλασιαστεί.

Ο νόμος αυτός μπορεί να ερμηνευτεί αν λάβουμε υπόψιν ότι το γλωσσικό σύστημα προσπαθεί να εξισορροπήσει τη συχνότητα μιας λέξης με τον αριθμό των λέξεων που μοιράζονται την ίδια συχνότητα εμφάνισης. Επίσης ο Wyllys (1981) περιγράφει στην εργασία του το νόμο του Zipf αναλυτικότερα. Ο Wyllys λοιπόν υποστηρίζει ότι ο νόμος αυτός περιγράφεται και ως:  $\log r \cdot \log f = \log c$ .

Επιπλέον για το  $\log ic$  ο Wyllys (1981) αναφέρει ότι ο υπολογισμός με βάση το νόμο του Zipf έχει πιο έγκυρα αποτελέσματα κυρίως όσον αφορά την κατάταξη λέξεων με μεσαία τάξη εμφάνισης παρά για τις λέξεις με πολύ υψηλή ή χαμηλή συχνότητα εμφάνισης.



**Εικόνα 2:** Συχνότητα εμφάνισης των λέξεων του ΘΕΓ σύμφωνα με τον πρώτο νόμο του Zipf

### 3.2. Νόμος των Menzerath-Altmann

Ο νόμος αυτός ανήκει ξεκίνησε από τον Ψυχολόγο-Φωνολόγο Menzerath ο οποίος θεώρησε ότι υπάρχει σχέση μεταξύ του συστατικού ενός ήχου και του συνόλου των ήχων μέσα στους οποίους βρίσκεται ο συγκεκριμένος ήχος. Έχει χαρακτηριστεί ως ένας από τους σημαντικότερους νόμους της ποσοτικής γλωσσολογίας και ουσιαστικά περιγράφει τη σχέση μεταξύ δομήματος (construct) και συστατικού (constituent), όπου όσο μεγαλώνει το πρώτο μικραίνει το δεύτερο.

Το 2013 ο Eroglu γράφει για τον Menzerath ότι είναι ο πρώτος που ανέδειξε την αρνητική συσχέτιση στη σχέση μεταξύ του μήκους ενός γλωσσικού δομήματος με τα συστατικά του. Λίγα χρόνια αργότερα ο Gabriel Altmann διατύπωσε με μαθηματικό τρόπο την παραδοχή του Menzerath. Οι δυο επιστήμονες εφάρμοσαν τον νόμο αυτό που είναι γνωστός ως Νόμος Menzerath-Altmann και διαπίστωσαν ότι μέσα σε μια πρόταση όταν ο αριθμός των συλλαβών αυξάνεται το μέσο μήκος της λέξης μειώνεται.

Προσπαθώντας να αναπαραστήσουμε μέσω μιας βασικής μαθηματικής σχέσης αυτή τη θεώρηση (νόμος Menzerath), το Z (ένας μέσος) μειώνεται καθώς το X αυξάνεται (όπου X ο αριθμός των τμημάτων της κατασκευής), ενώ το Y αντιπροσωπεύει το μέγεθος του συνόλου των τμημάτων, οπότε έχουμε (Buk & Rovennchak 2007):

$$Z = Y / X$$

Ο Altmann όμως διατύπωσε ένα άλλο μαθηματικό τύπο για αυτή την σχέση (Altman & Gerlach 2015):

$$y = aMx^b e^{-\gamma Mx}$$

όπου  $y$  είναι το του μέγεθος συστατικού (π. χ. μήκος συλλαβής),  $x$  — μέγεθος της γλωσσικής κατασκευής που θα ελεγχθεί (π. χ. αριθμός συλλαβών ανά λέξη),  $a, b, c$  είναι παράμετροι και  $e$  — ο αριθμός Euler. Ο αριθμός  $e$  είναι σημαντική μαθηματική σταθερά, η οποία αποτελεί τη βάση του φυσικού λογαρίθμου και προέρχεται από τον Ελβετό μαθηματικό Euler.

Μια προσέγγιση με την εφαρμογή του Νόμου Menzerath και Altmann πραγματοποιήθηκε από υποψήφιους διδάκτορες στο 7<sup>ο</sup> Συνέδριο Μεταπτυχιακών Φοιτητών & Υποψηφίων Διδασκτόρων του Τμήματος Φιλολογίας του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών. Εκεί λοιπόν παρουσιάστηκε ένα εργαλείο όπου έπαιρνε ένα σώμα κειμένου το οποίο αποτελούνταν από 736.000 λέξεις και στη συνέχεια γινόταν η κατηγοριοποίηση του περιεχομένου.

Στον παραπάνω πίνακα είναι ομαδοποιημένα τα μήκη των συλλαβών και των λέξεων μέσα στο σώμα κειμένων των 760.000 λέξεων και από αυτό το αποτέλεσμα μπορούμε να διαπιστώσουμε πως ισχύει η γραμμική εκδοχή του Νόμου των M-A καθώς το μέσο μήκος μια λέξης και μιας συλλαβής ανάλογα με το είδος του κειμένου δεν αλλάζει ιδιαίτερα. Επίσης παρατηρούμε ότι το μέσο μήκος της συλλαβής δεν μπορεί να έχει την ίδια τιμή με το μέσο μήκος των λέξεων αφού είπαμε ότι όταν αυξάνεται το ένα μειώνεται το άλλο.

Κατηγορίες	Κατηγορίες	Μέσος όρος (mean)	Μέγιστη τιμή (max)	Ελάχιστη τιμή (min)	Επιπρεπούσα τιμή (mode)	Διάμεσος (median)
Social Politics	MM Syl	2.3	2.6	2.2	2.3	2.3
	MM Word	2.4	3.2	1.9	2.5	2.4
Sports	MM Syl	2.4	2.8	2.1	2.3	2.3
	MM Word	2.3	3.2	1.8	2.4	2.3
Culture	MM Syl	2.4	2.7	2.2	2.3	2.4
	MM Word	2.3	2.8	1.8	2.0	2.3
Σύνολο (Whole corpus)	MM Syl	2.3	2.8	2.1	2.3	2.3
	MM Word	2.3	3.2	1.8	2.5	2.3

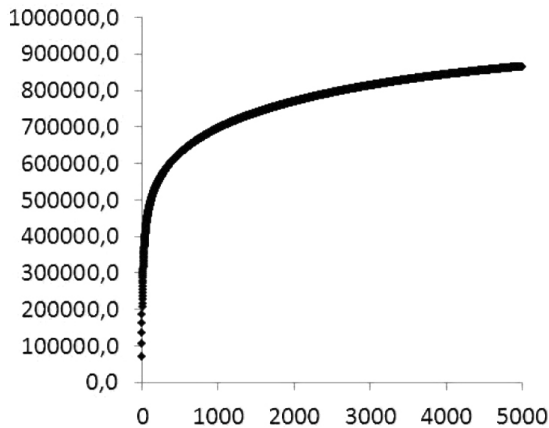
**Πίνακας 1:** Μέτρα κεντρικής τάσης μεταβλητών (MMSyl&MMWord) Corpus

### 3.3. Νόμος του Hearp

Μια άλλη χρήσιμη εκτίμηση που σχετίζεται με την συχνότητα εμφάνισης των λέξεων είναι η αύξηση του λεξιλογίου, ένα στοιχείο που συνδέεται άμεσα με το μέγεθος των ευρετηρίων σε ένα σύστημα ανάκτησης. Οι νέες λέξεις που προστίθενται σε μια συλλογή είναι λιγότερες όσο μεγαλύτερη είναι η συλλογή. Νέες λέξεις βέβαια πάντα θα υπάρχουν. Τέτοιες λέξεις είναι τα ονόματα, ανορθόγραφες λέξεις, αριθμοί, emails, κ. λπ.). Η σχέση μεταξύ του μεγέθους της συλλογής και του μεγέθους του λεξιλογίου ευρέθηκε εμπειρικά από τον Hearp το 1978 ότι ακολουθεί τον νόμο των δυνάμεων:

$$v = k \cdot n^\beta$$

όπου  $v$  είναι ο αριθμός των διακριτών λέξεων σε ένα κείμενο μεγέθους  $n$  και οι  $k$  &  $\beta$  είναι παράμετροι που προσδιορίζονται εμπειρικά. Σε ένα αγγλικό κείμενο για παράδειγμα το  $k$  κυμαίνεται μεταξύ 10 και 100 και το  $\beta$  μεταξύ 0.4 και 0.6. Για ένα βιβλίο η τιμή του παυξάνεται από την πρώτη λέξη ή συλλαβή μέχρι το τέλος της στο βιβλίο, έτσι ώστε το  $v(n)$  να σχηματίζει μια καμπύλη.



**Εικόνα 3:** Απεικόνιση του γραφήματος του Νόμου του Hearp

Στην παραπάνω εικόνα, στο γράφημα του Νόμου του Hearp ο άξονας  $x$  αντιπροσωπεύει το μέγεθος του κειμένου και ο άξονας  $y$  τον αριθμό των διακριτών στοιχείων λεξιλογίου που υπάρχουν στο κείμενο.

### 3.4. Ποσοτική γλωσσολογία και αυτόματη περίληψη κειμένου

Κατά την διαδικασία της εξαγωγής αυτόματης περίληψης θα πρέπει το σύστημα που θα εκτελέσει αυτή την διαδικασία να μπορεί να διαχωρίσει με κάποιον τρόπο τις λέξεις τις οποίες θα συμπεριλάβει στην περίληψή του. Αφού λοιπόν περάσει το πρώτο στάδιο της προεπεξεργασίας του κειμένου στην συνέχεια έρχεται η ποσοτική ανάλυση με τους νόμους που αναφέραμε παραπάνω να δώσει λύση στο πρόβλημα των λέξεων που πρόκειται να χρησιμοποιηθούν.

Ένα κείμενο το οποίο καλούμαστε να εξάγουμε την περίληψή του είναι ένα μέσο επικοινωνίας μεταξύ του των ανθρώπων. Σύμφωνα με τον Bolshakov & Gelbukh (2004) υπάρχουν τρία βασικά χαρακτηριστικά τα οποία διέπουν ένα κείμενο. Το πρώτο αφορά στον σκοπό ύπαρξης ενός κειμένου, δηλαδή στο ότι ένα κείμενο γεννιέται ώστε να κωδικοποιήσει ένα σύνολο πληροφοριών, επομένως έχει μια σημασία, ένα νόημα το οποίο προορίζεται και αφορά κάποιους ανθρώπους. Για το λόγο αυτό ακριβώς πραγματοποιείται η επεξεργασία της φυσικής γλώσσας του κειμένου. Το δεύτερο χαρακτηριστικό συνίσταται στο ότι όσες πληροφορίες και αν εμπεριέχονται σε ένα κείμενο, όσο πολύπλοκο και αν είναι η δομή του είναι πάντοτε γραμμική, αποτελεί δηλαδή ένα σύνολο από συμβολοσειρές, κενά, σημεία στίξης τα οποία σχηματίζουν μια πολύ μεγάλη γραμμή. Επίσης είναι σημαντικό να σημειωθεί ότι οι πληροφορίες που αναπαρίστανται στο κείμενο με γραμμική δομή είναι μη γραμμικές. Το τρίτο χαρακτηριστικό αφορά τη δομή του κειμένου από επιμέρους στοιχεία, τα οποία βρίσκονται «εντεθειμένα σε ομοειδή δομή» (nestedstructure) μέσα στο κείμενο.

Επιπλέον, ο Johnson όρισε τους βασικούς στόχους της ποσοτικής γλωσσολογίας: (α) μετατροπή δεδομένων (μείωση των ποσοτικών δεδομένων στα ουσιαστικά τους τμήματα ώστε να περιγραφούν τάσεις και κοινά στοιχεία), (β) διεξαγωγή συμπερασμάτων (μέσω ελέγχου υποθέσεων), (γ) έρευνα και περιγραφή σχέσεων μεταξύ των δεδομένων.

Μέσω της ποσοτικής ανάλυσης σύμφωνα με το (Mikros 2005) μπορούμε να κάνουμε ανάλυση της γλωσσικής χρήσης, δηλαδή να δημιουργήσουμε λίστες συχνότητας λέξεων ή λεξικά σύμπλοκα, να κατανοήσουμε από πού προέρχεται το κάθε κείμενο και να αναγνωρίσουμε το ύφος του συγγραφέα όποτε κάθε φορά που συναντάμε αυτό το ύφος γραφής να αναγνωρίζουμε αυτόματα και τον συγγραφέα. Άρα, έχοντας ένα κείμενο στα χέρια μας, μέσω της ποσοτικής ανάλυσης μπορούμε να πάρουμε ενδιαφέρουσες πληροφορίες τις οποίες θα συμπεριλάβουμε στην περίληψή μας, θα καθορίσουν την δομή της και θα μας βοηθήσουν να κατανοήσουμε το περιεχόμενο



της και σε ποιους απευθύνεται. Για παράδειγμα, ένα κείμενο έχει διαφορετική δομή που παρουσιάζει ένα επιστημονικό άρθρο που έχει μεγάλες προτάσεις με επιστημονικούς όρους και πλούσιο λεξιλόγιο και τελείως διαφορετικό είναι το κείμενο που συναντάμε σε ένα blog. Εκεί οι προτάσεις είναι απλές ή ελλειπτικές και πολύ πιο κατανοητές.

## 4. Ποσοτική ανάλυση της Τσακώνικης διαλέκτου

Στην μελέτη που ακολουθεί προσπαθήσαμε να εφαρμόσουμε ποσοτικές μεθόδους μέτρησης σε ένα λεξικό της Τσακώνικης διαλέκτου με σκοπό να εξαγουμε συμπεράσματα όσον αφορά το μέσο μήκος των λέξεων και να τα συγκρίνουμε με τη δημοτική, καθώς επίσης και να εξετάσουμε την συχνότητα εμφάνισης των γραμμάτων σε μια λέξη. Για την έρευνα μας χρησιμοποιήθηκε το λεξικό του Deffner το οποίο όπως έχει ήδη αναφερθεί είναι το πρώτο λεξικό που συναντάμε για τα Τσακωνικά και όπως αυτό βρέθηκε ψηφιοποιημένο από το δημοτικό σχολείο Λεωνιδίου.

Λεξικό τσακώνικης διαλέκτου	
Αρχική σελίδα	Εγγραφή
Σύνθεση	Αποσύνθεση
Οι λέξεις μου	
Α Β Γ Δ Ε Ζ Η Θ Ι Κ Λ Μ Ν Ξ Ο Π Ρ Σ Τ Υ Φ Χ Ψ α	
Λεξικό τσακώνικης διαλέκτου του Μιχαήλ Δέφνερ από τους μαθητές του Δημοτικού Σχολείου Λεωνιδίου	
α:	η
Σημειώσεις:	α μάτη=η μητέρα, η μάνα
α:	α
Σημειώσεις:	Επιφώνημα
α:	ας
Σημειώσεις:	α μόλη=ας έρθει
α:	αυ
Σημειώσεις:	πριν από σύμφωνο
άα:	άλλαξα
Σημειώσεις:	Αόριστος του ρήματος άσασου=αλλάζω
άβα (α):	άλλη (η)
αβαίνου:	λαμβάνω
Συνώνυμα:	λαβαίνου
Σημειώσεις:	αορ.: αβήκα μελ.:άβου
αβάκα:	κρυφά
Σημειώσεις:	Αβάκα επεράτσε α Τσουρακά

Εικόνα 4: Ψηφιοποιημένο λεξικό

Το περιεχόμενο του λεξικού μεταφέρθηκε σε ένα βιβλίο εργασίας Excel. Αυτό το βιβλίο εργασίας αποτελείται από 24 φύλλα. Το κάθε φύλλο περιλαμβάνει τις λέξεις από διαφορετικό γράμμα του αλφάβητου. Το κάθε φύλλο εργασίας αποτελείται από τρεις στήλες, στην πρώτη βρίσκεται η λέξη όπως αυτή αποτυπώνεται στα Τσακωνικά, η δεύτερη στήλη είναι η ερμηνεία της στη δημοτική και η τρίτη μας πληροφορεί ουσιαστικά σε ποια περιοχή της Τσακωνιάς συναντάμε τη λέξη αυτή.

738	αχουροχάραρε (το):	μεγαλο χαράρι (το)	
739	αχούχλιστε (ο):	ακόχλαστος (ο) , άβραστος	
740	αχρά (α):	αχλάδι (το)	
741	αχραΐα (α):	αχλαδιά (η)	
742	αχραούνα (α):	αχλαδάρα (η)	
743	αχραούνι (το):	αχλαδάκι (το)	
744	αχραόφυλλε (το):	αχλαδόφυλλο (το)	
745	αχρείε, -α, -ε (ο):	αχρείος (ο) , άθλιος , ελεεινός	
746	αχρία (α):	τσουκνίδα (η)	
747	αχρόνιστε (ο):	αχρόνιστος (ο)	
748	αχρείαστε (ο):	αχρείαστος (ο)	
749	άχρηστε (ο):	άχρηστος (ο)	
750	αχτάρα (α):	λαχτάρα (η)	
751	αχταριστέ, -ά, -ε (ο):	λαχταριστός (ο)	
752	αχτένιστε (ο):	αχτένιστος (ο)	
753	αχυρζάνα (ο):	αχυρώνας (ο)	
754	αχωμάκιστε (ο):	αχωμάτιστος (ο)	
755	αχώνευτε (ο):	αχώνευτος (ο)	
756	άφατε (ο):	ασκούπιστος (ο)	
757	άφαητε (ο):	άφαχος (ο) , αφηλάφητος	
758	αφεά:	ψηλά , με δυνατή φωνή , δυνατά	
759	αφελέ (ο):	ψηλός (ο)	
760	αψηφού:	αψηφώ , αμελώ , αδιαφορώ	απεινίδι
761	αψηφούκουντά ένι:	βραδιάζει	ο
762	αψίφουμα (το):	σκοτείνιασμα (το) , σουρούπημα	
763	άψυχο (ο):	άψυχος (ο) , αδύνατος	απεινίδι ο
764	αψόφητε (ο):	αψόφητος (ο) , ζωντανός	απεινίδι ο
765	αψώνιστε (ο):	αψώνιστος (ο)	απεινίδι ο
766	άωνα (α):	αλώνι (το)	
767	σημειώσεις:	στην Καστάνιτσα άλωνα	
768	αωνάρι (α):	Ιούλιος (ο) , αλωνάρης	
769	αωνάτζι (το):	αλωνάκι (το)	
770	αωνού:	αλωνίζω	

**Εικόνα 5:** Το λεξικό καταχωρημένο σε βιβλίο εργασίας Excel

Για την υλοποίηση της μέτρησης χρησιμοποιήθηκε η γλώσσα προγραμματισμού Python μέσω του Jupiter Notebook. Το Jupiter Notebook είναι μια διαδικτυακή εφαρμογή ελεύθερου λογισμικού και ανοικτού κώδικα το οποίο δίνει τη δυνατότητα δημιουργίας και διαμοιρασμού κώδικα για τον υπολογισμό εξισώσεων, πιθανοτήτων και στατιστικών στοιχείων. Υποστηρίζεται επίσης από 40 γλώσσες προγραμματισμού. Επειδή θέλουμε να ασχοληθούμε για χειρισμό και ανάλυση δεδομένων χρησιμοποιούμε την βιβλιοθήκη λογισμικού “pandas”. Επίσης εισάγεται και η βιβλιοθήκη “re” για τον χειρισμό των κανονικών εκφράσεων η οποία είναι χρήσιμη για την εισαγωγή των δεδομένων από το βιβλίο excel.

```
import pandas as pd
import re
df=pd.read_excel("leksiko.xlsx",sheet_name="A",header=None).rename
({0:'word',1:'description',2:'location'},axis=1).dropna(how="all")
```

Όπως έχουμε ήδη αναφερθεί το λεξικό που είναι καταχωρημένο στο excel είναι χωρισμένο σε φύλλα εργασίας. Με την παραπάνω εντολή διαβάζουμε το συγκεκριμένο λεξικό και κάθε φορά που θέλουμε να διαβάσουμε διαφορετικό φύλο εργασίας απλά αλλάζουμε την τιμή στην μεταβλητή «sheet name».

	word	description	location
<b>count</b>	620	620	248
<b>unique</b>	562	601	6
<b>top</b>	Συνώνυμα:	ασκέπαστος (ο)	Λεωνίδιο
<b>freq</b>	20	2	148

**Εικόνα 6:** Το σύνολο των λέξεων στο φύλλο «A»

Θέλουμε να υπολογίσουμε το μέσο μήκος των λέξεων για τις λέξεις που ξεκινούν με «A». Χρησιμοποιούμε την παρακάτω εντολή. Αξίζει να σημειωθεί ότι στο μέσο μήκος της λέξης δεν υπολογίζονται τα κενά που τυχόν υπάρχουν, ούτε σημεία στίξης ή ειδικοί χαρακτήρες.

```
df['wordLen']=df['word'].apply(lambdax:len(x.replace(",").replace("(",").replace(")",").replace(":","))-1 if '(' in x else len(x))
```

	word	description	location	wordLen
<b>1</b>	αγό (ο):	λαγός (ο)	Λεωνίδιο	3
<b>4</b>	άγο του θεού (το):	άλογο του θεού (το)	Λεωνίδιο	11
<b>9</b>	αγόμυλε (ο):	αλογόμυλος (ο)	NaN	7
<b>11</b>	αγοράνου:	αγοράζω, ψωνίζω	Μέλινα	10
<b>13</b>	αγούμαστε (ο):	αγούμαστος (ο), είδος σταφυλιού	Μέλινα	9
...	...	...	...	...
<b>766</b>	άωνα (ο):	αλώνι (το)	NaN	5
<b>768</b>	αωνάρι (το):	Ιούλιος (ο), αλωνάρης	NaN	7
<b>769</b>	αωνάτζι (το):	αλωνάκι (το)	NaN	9
<b>770</b>	αωνού:	αλωνίζω	NaN	7
<b>771</b>	άωρα:	αργά τη νύχτα	NaN	6

559 rows x 4 columns

**Εικόνα 7:** Μήκος λέξεων στο φύλλο «A»

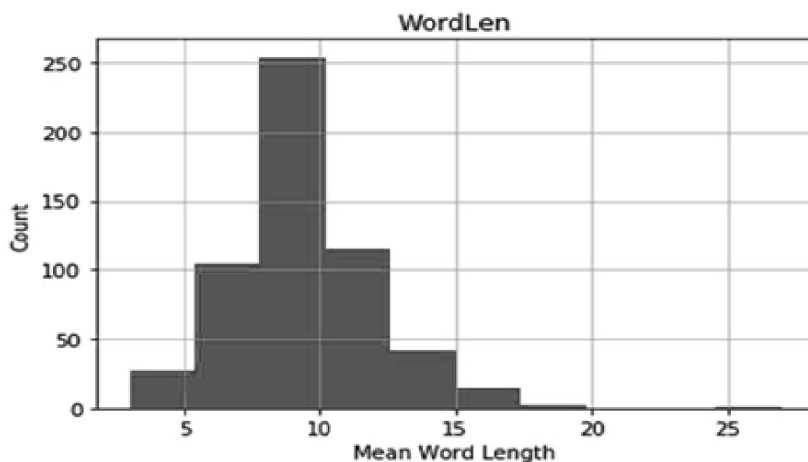
Το αποτέλεσμα της παραπάνω εντολής μας επιστρέφει ένα πίνακα όπου απεικονίζεται σε μια ξεχωριστή στήλη το μήκος της λέξης. Επιπλέον μπορούμε να δούμε το ποσοστιαίο μήκος της λέξης όπως περιγράφεται στον παρακάτω πίνακα όπου παρουσιάζει τη μέση τιμή και την τυπική απόκλιση των αποτελεσμάτων.

```
In [248... df.describe ()
```

wordLen	
count	559.000000
mean	9.406082
std	2.588512
min	3.000000
25%	8.000000
50%	9.000000
75%	11.000000
max	27.000000

**Εικόνα 8:** Μήκος λέξης σε ποσοστά

Παρακάτω παρατηρούμε ότι από ένα σύνολο 559.000 χαρακτήρων, το μέσο μήκος μιας λέξης που αρχίζει με «α» είναι το 9,4 και η τυπική της απόκλιση 2,4. Αυτό σημαίνει ότι μια λέξη που αρχίζει με αυτό το γράμμα έχει ένα μέσο αριθμό χαρακτήρων 9, ενώ βασίζόμενοι στην τυπική απόκλιση παρατηρούμε ότι είναι χαμηλή. Άρα δεν υπάρχει μεγάλη διασπορά.



**Εικόνα 9:** Διαγραμματική απεικόνιση των αποτελεσμάτων μέσω μήκος λέξεων σε συνάρτηση με το πλήθος

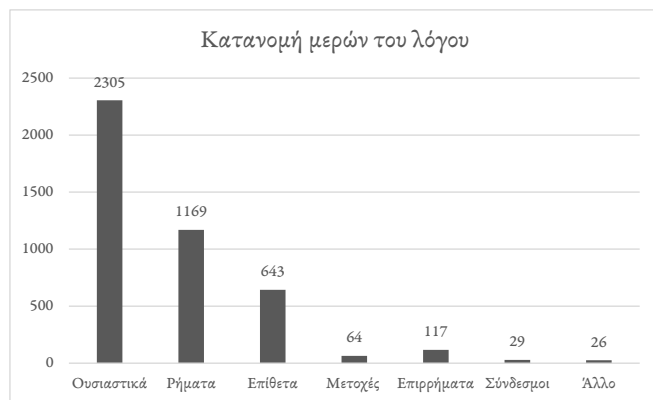
Με τον ίδιο τρόπο έγινε η υλοποίηση για όλα τα γράμματα από το Α μέχρι το Ω. Και παρατηρήσαμε ότι δεν υπήρχαν μεγάλες αποκλίσεις των τιμών μεταξύ τους. Επιπλέον, πραγματοποιήθηκε και σύγκριση μεταξύ των λέξεων της Τσακωνικής και της Δημοτικής.

```
df['TS_words']=df.Tsakonika.apply(get_words)
df['Tsakonika_len']=df.TS_words.apply(len)
df['DHM_words']=df.Dhmotikh.apply(get_words)
df['DHM_len']=df.DHM_words.apply(len)
```

	Τσακωνικά	Ελληνικά	Location	Τσακωνικά_len	TS_words	DHM_words	DHM_len
0	αγό (α)	Αγορά (α)	Λευκωία	3	αγο	Αγορά	5
3	αγο τσε φουα (α)	άγο τσε φουά (α)	Λευκωία	3	αγο	άγο	5
10	αγοτσίμ	αγοτσίμ, φουτσίμ	Μόνα	8	αγοτσίμ	αγοτσίμ	7
12	αγομπίσι (α)	αγομπίσι (α), είδος στεφάνου	Μόνα	9	αγομπίσι	αγομπίσι	10
14	αγομρέ + ι (α)	αγομρέ (α)	Λευκωία	6	αγομρέ	αγομρέ	7
17	αγομμεντέφ,αγομμεντένο (α)	αγομμεν (α)	Λευκωία	13	αγομμεντέφ	αγομμεντένο	10
21	αγομπίσι + ι (α)	αγομπίσι (α)	Λευκωία	9	αγομπίσι	αγομπίσι	10
26	αγομπίσι (α)	αγομπίσι (α)	Μόνα	8	αγομπίσι	αγομπίσι	11
29	αγομ + ι (α)	αγομ (α), άγομ (α)	Λευκωία	4	αγομ	άγομ	6
32	αγομπίσι	αγομπίσι, έγομπίσι, έγομπίσιμ	Λευκωία	8	αγομπίσι	αγομπίσι	7
37	αγομ (α)	αγομ (α), άγομ (α)	Λευκωία	6	αγομ	αγομ	6
40	αγομπίσι (α)	αγομ πούσι (α)	Λευκωία	11	αγομπίσι	αγομ	5
44	αγομπίσι (α)	αγομπίσι (α)	Μόνα	12	αγομπίσι	αγομπίσι	9
49	αγομπίσι	αγομπίσι, απομπίσι, απομπίσιμ	Λευκωία	4	αγομ	αγομπίσι	10
54	αγομπίσι	αγομπίσι, απομπίσι, απομπίσιμ	Μόνα	8	αγομπίσι	αγομπίσι	7
54	αγομπίσι (α)	απομπίσι (α), άπομπίσι (α)	Λευκωία	4	αγομ	απομπίσι	9
62	αγομπίσι (α)	αγομπίσι (α)	Λευκωία	12	αγομπίσι	αγομπίσι	10
66	αγομπίσι (α)	αγομπίσι (α)	Λευκωία	11	αγομπίσι	αγομπίσι	12
69	αγομπίσι (α)	αγομ αλίσι (α)	Λευκωία	9	αγομπίσι	αγομ	5
73	αγομπίσι (α)	αγομπίσι (α)	Λευκωία	7	αγομπίσι	αγομπίσι	8
75	αγομπίσι (α)	αγομπίσι (α)	Μόνα	12	αγομπίσι	αγομ	5
77	αγομπίσι (α)	αγομπίσι (α)	Λευκωία	7	αγομπίσι	αγομπίσι	10
80	αγομπίσι (α)	αγομπίσι (α)	Λευκωία	8	αγομπίσι	αγομπίσι	8
89	άδοσι (α)	άδοσι (α), αδοσι (α)	Μόνα	6	άδοσι	άδοσι	7
93	άδοσι (α)	άδοσι (α)	Λευκωία	5	άδοσι	άδοσι	8
96	άδοσι (α)	άδοσι (α), άδοσι (α)	Μόνα	8	άδοσι	άδοσι	6
98	άδοσι (α)	άδοσι (α), άδοσι (α)	Μόνα	5	άδοσι	άδοσι	6
102	άδοσι + ι (α)	άδοσι (α), άδοσι (α)	Λευκωία	6	άδοσι	άδοσι	6
118	άδοσι + ι (α)	άδοσι (α), άδοσι (α)	Μόνα	5	άδοσι	άδοσι	7
112	άδοσι (α)	άδοσι (α), άδοσι (α)	Λευκωία	5	άδοσι	άδοσι	8
--	--	--	--	--	--	--	--
609	αλόγι (α)	αλόγι (α)	Λευκωία	8	αλόγι	αλόγι	9
613	άνοσι (α)	άνοσι (α)	Λευκωία	5	άνοσι	άνοσι	6
634	απομπίσι (α)	απομπίσι (α)	Λευκωία	8	απομπίσι	απομπίσι	8
636	απομπίσι (α)	απομπίσι (α)	Λευκωία	9	απομπίσι	απομπίσι	8
637	απομπίσι (α)	απομπίσι (α)	Λευκωία	6	απομπίσι	απομπίσι	7
640	απομπίσι	απομπίσι	Λευκωία	10	απομπίσι	απομπίσι	9
641	απομπίσι	απομπίσι	Λευκωία	6	απομπίσι	απομπίσι	5
647	απομπίσι (α)	απομπίσι (α)	Λευκωία	7	απομπίσι	απομπίσι	8
664	απομπίσι (α)	απομπίσι (α)	Λευκωία	7	απομπίσι	απομπίσι	8
685	απομπίσι (α)	απομπίσι (α)	Λευκωία	5	απομπίσι	απομπίσι	6
687	απομπίσι (α)	απομπίσι (α)	Λευκωία	8	απομπίσι	απομπίσι	9
688	απομπίσι (α)	απομπίσι (α)	Λευκωία	9	απομπίσι	απομπίσι	9
690	απομπίσι (α)	απομπίσι (α)	Λευκωία	9	απομπίσι	απομπίσι	10
694	απομπίσι (α)	απομπίσι (α)	Λευκωία	11	απομπίσι	απομπίσι	12
695	απομπίσι (α)	απομπίσι (α)	Λευκωία	3	απομπίσι	απομπίσι	9
699	απομπίσι (α)	άδοσι (α), άδοσι (α)	Λευκωία	6	απομπίσι	άδοσι	8
700	απομπίσι (α)	άδοσι (α)	Λευκωία	7	απομπίσι	άδοσι	8
701	απομπίσι (α)	άδοσι (α)	Λευκωία	10	απομπίσι	άδοσι	9
708	απομπίσι (α)	απομπίσι (α)	Λευκωία	10	απομπίσι	απομπίσι	10
709	απομπίσι	απομπίσι	Λευκωία	7	απομπίσι	απομπίσι	4
711	απομπίσι (α)	απομπίσι (α)	Λευκωία	5	απομπίσι	απομπίσι	6
712	απομπίσι (α)	απομπίσι (α)	Λευκωία	5	απομπίσι	απομπίσι	9
714	απομπίσι (α)	απομπίσι (α)	Λευκωία	8	απομπίσι	απομπίσι	9
715	απομπίσι (α)	απομπίσι (α)	Λευκωία	5	απομπίσι	απομπίσι	6
718	απομπίσι (α)	απομπίσι (α)	Λευκωία	9	απομπίσι	απομπίσι	6
722	απομπίσι (α)	απομπίσι (α)	Λευκωία	10	απομπίσι	απομπίσι	9
724	απομπίσι (α)	απομπίσι (α)	Λευκωία	7	απομπίσι	απομπίσι	2
729	απομπίσι	απομπίσι, απομπίσι, απομπίσιμ	Λευκωία	6	απομπίσι	απομπίσι	5
762	απομπίσι (α)	απομπίσι (α), άδοσι (α)	Λευκωία	6	απομπίσι	άδοσι	6
763	απομπίσι (α)	απομπίσι (α), άδοσι (α)	Λευκωία	7	απομπίσι	άδοσι	8

Εικόνα 10: Μήκος λέξεων Τσακωνικής και Δημοτικής

Ολοκληρώνοντας, πραγματοποιήθηκε η κατανομή μερών του λόγου και η διαγραμματική τους απεικόνιση φαίνεται παρακάτω.



**Εικόνα 11.** Διαγραμματική απεικόνιση των μερών του λόγου

## Συμπεράσματα

Από την έρευνα που πραγματοποιήθηκε μελετώντας το λεξικό του Deffner και εφαρμόζοντας τις μεθόδους γλωσσολογικής έρευνας καταλήγουμε στο συμπέρασμα ότι το μεγαλύτερο μέρος του λεξικού αυτού της Τσακωνικής διαλέκτου αποτελείται από ουσιαστικά, ρήματα και επίθετα. Τα ουσιαστικά που είναι καταγεγραμμένα ανέρχονται στα 2305, ενώ τα μέρη του λόγου όπως μετοχές, επιρρήματα σύνδεσμοι και άλλα είναι πολύ λιγότερα. Από αυτό μπορούμε να συμπεράνουμε ότι ο λόγος τους ήταν απλός και κατανοητός στην μεταξύ τους επικοινωνία παρά το γεγονός ότι θεωρείται μια δυσνόητη γλώσσα.

Επιπλέον, από τις μετρήσεις που κάναμε για το μήκος των λέξεων και, παρατηρούμε ότι το μήκος μιας λέξης κυμαίνεται από 3 το ελάχιστο έως 20 χαρακτήρες το μέγιστο. Στο μεγαλύτερο ποσοστό το ελάχιστο μήκος λέξεων είναι οι 7 χαρακτήρες και το μέγιστο οι 10. Ενώ, η μέγιστη συχνότητα εμφάνισης ενός γράμματος είναι το 9,4. Από αυτό μπορούμε να κατανοήσουμε την δυσκολία που αντιμετωπίζουν ως προς την ανάγνωση και την κατανόηση αυτής της διαλέκτου, λόγω των μεγάλων λέξεων και της δύσκολης εκφοράς τους.

Για περαιτέρω διερεύνηση τίθεται το θέμα του αριθμού των συμφώνων και των φωνηέντων σε μια λέξη. Καθώς επίσης και η μελέτη και διερεύνηση του λεξικού του Κώστακη

το οποίο αποτελείται από περισσότερες καταγεγραμμένες λέξεις αλλά και η σύγκριση με αυτό του Deffner.

## Ευχαριστίες

Με την περάτωση του άρθρου αυτού μου δίνεται η ευκαιρία να ευχαριστήσω την κ. Σοφία Στάμου επιβλέπουσα καθηγήτρια στη διδακτορική μου διατριβή και συνεργάτη στη συγγραφή αυτού του άρθρου.

Θα πρέπει παράλληλα να ευχαριστήσουμε τους διοργανωτές αυτού του Συνεδρίου κ. Αγγελική Ράλλη, κ. Brian Joseph, κ. Mark Janse & κ. Maxim Kisilier. Ένα μεγάλο ευχαριστώ, ιδιαίτερα στον κ. Maxim Kisilier για την υποστήριξη και βοήθεια του και τον κ. Παναγιώτη Ιωακειμίδα για την συμπαράστασή του.

Τέλος, ένα μεγάλο ευχαριστώ στις οικογένειές μας.

## Βιβλιογραφία

- Altman, E. G. & M. Gerlach 2015. Statistical law in linguistics. In M. Degli Esposti, E. G. Altmann & F. Pachet (eds.), *Creativity and Universality in Language*. Paris: Springer Cham, 7–26.
- Bolshakov, I. & A. Gelbukh. 2004. A very large dictionary with paradigmatic, syntagmatic, and paronymic. In M. Zock & P. Saint Disier (eds.), *Proceedings of the Workshop on Enhancing and Using Electronic Dictionaries*. Geneva: COLING, 53–56.
- Buk, S. & A. Rovennchak 2007. Menzerath-Altamann Law for Syntactic Structures in Ukrainian. Available at: <https://arxiv.org/pdf/cs/0701194.pdf> (accessed on 19.02.2013).
- Deffner, M. 1881. *Zakonishe Grammatik*. Bd. I: Lautlehre. Berlin: Weidmannsche Buchhandlung.
- Hatzidakis, G. 1892. *Einleitung in die neugriechische Grammatik*. Leipzig: Druck und Verlag von Breitkopf & Härtel.
- Mikros, G. 2005. Quantitative linguistics in Greece. In R. Köhler, G. Altmann & R. G. Piotrowski (Hrsg.), *Quantitative Linguistik — Quantitative Linguistics. An International Handbook*. Berlin & New York: De Gruyter, 136–142.
- Wyllys, R. E. 1981. Empirical and theoretical bases of Zipf's law. *Library Trends* 30(1): 53–64.
- Δέφνερ, Μ. 1923. *Λεξικόν τῆς Τσακωνικῆς διαλέκτου*. Αθήνα: Εστία.
- Κοντοσόπουλος, Ν. 2001. *Διάλεκτοι και ιδιώματα της Νέας Ελληνικής*. Αθήνα: Εκδόσεις Γρηγόρη.
- Κωστάκης, Θ. Π. 1986α. *Λεξικό της Τσακωνικής διαλέκτου*. Τ. Α'. Αθήνα: Ακαδημία Αθηνών.
- Κωστάκης, Θ. Π. 1986β. *Λεξικό της Τσακωνικής διαλέκτου*. Τ. Β'. Αθήνα: Ακαδημία Αθηνών.
- Κωστάκης, Θ. Π. 1987. *Λεξικό της Τσακωνικής διαλέκτου*. Τ. Γ'. Αθήνα: Ακαδημία Αθηνών.
- Μηνάς, Κ. 2004. Φωνητικά ισόγλωσσα Τσακωνίας και Μάνης. Στο Κ. Μηνάς, *Μελέτες Νεοελληνικής διαλεκτολογίας*. Αθήνα: Τυπωθήτω & Γιώργος Δάρδανος, 534–551.



- Μικρός, Γ., Ν. Χατζηγεωργίου & Γ. Καραγιάννης. 2003. Βασικά ποσοτικά μεγέθη στην γραπτή Νέα Ελληνική γλώσσα: η αξιοποίηση του ΕΘΕΓ στην ελληνική ποσοτική γλωσσολογία. In *Proceedings of the Workshop Text Processing for Modern Greek: From Symbolic to Statistical Approaches*. Rethymno: University of Cyprus, 23–37.
- Οικονόμου, Θ. 1870. *Γραμματική τῆς τσακωνικῆς διαλέκτου*. Αθήνα: Εκ του τυπογραφείου Λ. Ψυλλιάκου και Σ<sup>α</sup>.
- Τζιτζιλής, Χ. 2000. *Νεοελληνικές διάλεκτοι και νεοελληνική διαλεκτολογία*. Στο Α. Φ. Χριστίδης, Μ. Αραποπούλου & Γ. Γιαννουλοπούλου (επίμ.), *Η ελληνική γλώσσα και οι διάλεκτοί της*. Αθήνα: ΥΠΕΠΘ & Κέντρο Ελληνικής Γλώσσας, 15–22.
- Χατζιδάκις, Γ. 1905. *Μεσαιωνικά και Νέα Έλληνικά*. Τ. Α'. Αθήνα: Τύποις Π. Δ. Σακελλαρίου.