



Τεχνολογικό
Πανεπιστήμιο
Κύπρου

Σχολή Επιστημών Υγείας

Πτυχιακή εργασία

**ΚΑΤΑΜΕΤΡΗΣΗ ΣΥΧΝΟΤΗΤΑΣ ΛΕΞΕΩΝ ΣΤΗ
ΚΥΠΡΙΑΚΗ ΓΛΩΣΣΑ**

Ηλιάνα Φράγκου

Λεμεσός, Μάιος 2020

ΤΕΧΝΟΛΟΓΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΑΠΟΚΑΤΑΣΤΑΣΗΣ

Πτυχιακή εργασία

ΚΑΤΑΜΕΤΡΗΣΗ ΣΥΧΝΟΤΗΤΑΣ ΛΕΞΕΩΝ ΣΤΗ ΚΥΠΡΙΑΚΗ

ΓΛΩΣΣΑ

της

Ηλιάνας Φράγκου

Επιβλέπων Καθηγητής

Δρ. Νίκος Κωνσταντίνου

Λεμεσός, Μάιος 2020

Πνευματικά δικαιώματα

Copyright © Ηλιάνα Φράγκου, 2020

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της πτυχιακής εργασίας από το Τμήμα Επιστημών Αποκατάστασης του Τεχνολογικού Πανεπιστημίου Κύπρου δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Τμήματος.

Θα ήθελα να ευχαριστήσω ιδιαίτερα την κ. Άρτεμις Τραϊκάπη, για τη συμβολή και την καθοδήγηση της στη συγγραφή της πτυχιακής εργασίας.

ΠΕΡΙΛΗΨΗ

Σκοπός Το «word frequency effect» αφορά τη συχνότητα στην οποία μια λέξη συναντιέται σε μια γλώσσα. Η συχνότητα των λέξεων μπορεί να χρησιμοποιηθεί αποτελεσματικά σε κλινικούς αλλά και σε μη κλινικούς πληθυσμούς. Στόχος της παρούσας μελέτης είναι η καταμέτρηση της συχνότητας στην οποία εμφανίζονται οι λέξεις στην Κυπριακή διάλεκτο, ώστε να δημιουργηθεί η πρώτη λίστα συχνότητας λέξεων σημασιολογικού περιεχομένου στον Κυπριακό πληθυσμό.

Μεθοδολογία Δέκα γνωστικά υγιείς συμμετέχοντες, Κυπριακής καταγωγής, με επίδοση άνω των 27 στο MMSE, επιλέχθηκαν να συμμετάσχουν στην παρούσα μελέτη. Στους συμμετέχοντες χορηγήθηκε η «Λίστα Συχνότητας Λέξεων» η οποία δημιουργήθηκε από την ομάδα της έρευνας. Η λίστα αποτελείται από δώδεκα κύριες σημασιολογικές κατηγορίες και ορισμένες υποκατηγορίες.

Αποτελέσματα Ανιχνεύθηκαν συνολικά 521 λέξεις. Από αυτές οι 374 λέξεις ήταν λέξεις χαμηλής συχνότητας, 81 λέξεις μέτριας συχνότητας και 66 λέξεις υψηλής συχνότητας. Στις σημασιολογικές κατηγορίες «Αντικείμενα» (126) και «Ζώα» (123) καταγράφηκαν οι περισσότερες λέξεις ενώ στις σημασιολογικές κατηγορίες «Ρούχα» (35) και «Λαχανικά» (31) καταγράφηκαν οι λιγότερες. Σε όλες τις σημασιολογικές κατηγορίες καταγράφηκαν περισσότερες λέξεις χαμηλής συχνότητας.

Συμπέρασμα Παρατηρήθηκαν κάποιες διαφορές ανάμεσα στην συχνότητα εμφάνισης των λέξεων στην Ελληνική γλώσσα και την Κυπριακή διάλεκτο. Καταγράφηκαν πολύ περισσότερες λέξεις χαμηλής παρά μέτριας και υψηλής συχνότητας. Η παρούσα λίστα μπορεί να χρησιμοποιηθεί στην κλινική πράξη για τη δημιουργία γνωστικών τεστ αξιολόγησης και την αποκατάσταση κλινικών πληθυσμών.

Λέξεις κλειδιά: συχνότητα, λέξεις, word frequency effect, λίστα συχνότητας, σημασιολογικές κατηγορίες

ABSTRACT

Objectives The «word frequency effect» relates to the frequency of a word appearance in a language. The frequency of words can be used efficiently both in clinical and non – clinical population. The aim of this study is to encounter the frequency at which words appear in Cypriot dialect, in order to create the first semantic words frequency list in the Cypriot population.

Methods Ten cognitively healthy participants, with Cypriot origin and a score above 27 at MMSE were chosen to participate in the current study. To the participants, was given a word frequency list which was made by the research team. That list consists of twelve main semantic categories and some subcategories.

Result A total of 521 words were detected. From them 374 words were low frequency words, 81 medium – frequency words and 66 were high – frequency words. The most words were recorded at the semantic categories “Objects” (126) and “Animals” (123) while the less were recorded at the semantic categories “Clothes” (35) and “Vegetables” (31). In all of the semantic categories more low – frequency words were encountered.

Conclusion Some differences were observed between the word’s frequency in Greek language and the Cypriot dialect. Based on all the data, much more low – frequency words were recorded than medium or high frequency. Current frequency list can be used in clinical practice in order to make cognitive tests and help clinical people with their recover.

Keywords: frequency, words, word frequency effect, frequency list, semantic categories

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ.....	vi
ABSTRACT.....	vii
ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ	viii
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ	x
ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΩΝ	xi
ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ	xii
ΑΠΟΔΟΣΗ ΟΡΩΝ	xiii
1 Εισαγωγή	1
1.1 “Word Frequency Effect”	1
1.2 Λέξεις υψηλής και χαμηλής συχνότητας	1
1.3 Μέτρηση συχνότητας.....	2
1.4 Σύγκριση λέξεων υψηλής συχνότητας έναντι λέξεων χαμηλής συχνότητας.....	2
1.5 Βάσεις δεδομένων αναφορικά με τη συχνότητα εμφάνισης λέξεων	3
1.6 Βάσεις δεδομένων συχνότητας στην Ελληνική γλώσσα	4
1.7 Χρησιμότητα συχνότητας	5
1.7.1 Μη κλινικοί σκοποί.....	5
1.7.2 Κλινικοί σκοποί	6
1.8 Διαφορές μεταξύ της Νέας Ελληνικής γλώσσας και της Κυπριακής διαλέκτου 10	
1.9 Σκοπός της παρούσας έρευνας.....	11
2 Μεθοδολογία Έρευνας.....	12
2.1 Δείγμα και κριτήρια εισαγωγής στην έρευνα.....	12
2.2 Διαδικασία συλλογής δεδομένων.....	12
2.3 Μέσα και τρόπος συλλογής δεδομένων	13

2.3.1	Mini – Mental State Examination – MMSE	13
2.3.2	Κύριο ερευνητικό εργαλείο	14
3	Αποτελέσματα.....	16
3.1	Περιγραφή δείγματος.....	16
3.2	Περιγραφικά Ευρήματα	16
3.3	Συχνότητα λέξεων ανά σημασιολογική κατηγορία.....	17
3.3.1	Υψηλής Συχνότητας Λέξεις.....	17
3.3.2	Μέτριας Συχνότητας Λέξεις	17
3.3.3	Χαμηλής Συχνότητας Λέξεις.....	18
4	Συζήτηση	19
5	Συμπεράσματα	21
	ΕΠΙΛΟΓΟΣ	23
	ΠΑΡΑΡΤΗΜΑ Ι.....	31
	ΠΑΡΑΡΤΗΜΑ ΙΙ.....	39

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1: Συχνότητα λέξεων σημασιολογικής κατηγορίας Ζώα	39
Πίνακας 2: Συχνότητα λέξεων σημασιολογικής κατηγορίας Υπεραγορά.....	41
Πίνακας 3: Συχνότητα λέξεων σημασιολογικής κατηγορίας Φαγητά.....	43
Πίνακας 4: Συχνότητα λέξεων σημασιολογικής κατηγορίας Λαχανικά.....	46
Πίνακας 5: Συχνότητα λέξεων σημασιολογικής κατηγορίας Ρούχα	47
Πίνακας 6: Συχνότητα λέξεων σημασιολογικής κατηγορίας Αντικείμενα.....	48

ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 1: Βαθμολογίες συμμετεχόντων στο γνωστικό εργαλείο MMSE..... 16

ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

WFE:	Word Frequency Effect
HNC:	Hellenic National Corpus
MMSE:	Mini – Mental State Examination
ADAS – Cog:	Alzheimer's Disease Assessment Scale-Cognitive Subscale

ΑΠΟΔΟΣΗ ΟΡΩΝ

Word Frequency	Συχνότητα Λέξεων
Low – frequency words	Χαμηλής συχνότητας λέξεις
High – frequency words	Υψηλής συχνότητας λέξεις

1 Εισαγωγή

1.1 “Word Frequency Effect”

Ο όρος “word frequency effect” (WFE), αναφέρεται στη συχνότητα στην οποία μια λέξη συναντιέται σε μια γλώσσα (Baayen, Feldman, & Schreuder, 2006). Έχει μια δυνατή επιρροή στην ικανότητα ενός ατόμου να ονομάζει, να αναγνωρίζει και να ανακαλεί μια λέξη. (Baayenet al., 2006). Η συχνότητα στην οποία εντοπίζονται οι λέξεις είναι σημαντικός παράγοντας για την απόδοση της μνήμης (Brysbart, Mandera & Keuleers, 2018). Το WFE μπορεί να έχει ξεχωριστή μορφή μεταξύ διάφορων ομάδων ανθρώπων, όπως για παράδειγμα σε άτομα με σχιζοφρένεια (Brébion, David, Bressan, & Pilowsky, 2005), με αμνησία (Ellis, Miller & Sin, 1983), Νόσο Αλτσχάιμερ (Balota, Burgess, Cortese & Adams, 2002) και με εγκεφαλικό επεισόδιο (Hoffman, Rogers & Lambon Ralph, 2011). Επίσης, σύμφωνα με τους Howes και Solomon (1951) (as cited in Brysbart & New, 2009) η συχνότητα των λέξεων είναι μια πολύ σημαντική μεταβλητή στη γνωστική επεξεργασία. Σύμφωνα με τον de Groot (2006), η συχνότητα των λέξεων σχετίζεται άμεσα με το πόσο γνωστές είναι οι έννοιες μιας γλώσσας στους ομιλητές της. Λέξεις οι οποίες χρησιμοποιούνται συχνότερα από τους ομιλητές μιας γλώσσας, είτε γραπτώς είτε στον προφορικό λόγο, φαίνεται πως αναφέρονται σε έννοιες οι οποίες είναι πιο γνωστές στους ομιλητές από ότι σε έννοιες που είναι δυσκολότερο να συναντηθούν σε μια γλώσσα (de Groot, 2006).

1.2 Λέξεις υψηλής και χαμηλής συχνότητας

Μια κύρια μεταβλητή που φανερώνει την σημαντικότητα του WFE, είναι η συχνότητα, βάσει της οποίας φαίνεται να υπάρχουν δύο κύριες κατηγορίες λέξεων (Badham et al., 2017). Οι λέξεις υψηλής συχνότητας (high – frequency words) και οι λέξεις χαμηλής συχνότητας (low – frequency words) (Brysbart et al., 2018). Οι υψηλής συχνότητας λέξεις αφορούν λέξεις οι οποίες παρουσιάζονται συχνότερα σε μια γλώσσα (Brysbart et al., 2018). Είναι γνωστές στους περισσότερους ομιλητές μιας γλώσσας (Monsell, Doyle & Haggard, 1989), επομένως επεξεργάζονται πιο γρήγορα σε περιπτώσεις κατονομασίας λέξεων, σημασιολογικής απόφασης ή όταν πρόκειται για λεξικές αποφάσεις (Brysbart, Keuleers, & New, 2011). Η έννοια λεξική απόφαση (lexical decision), αφορά τον χρόνο που απαιτείται ώστε να αποφασιστεί αν μια σειρά

γραμμάτων αποτελεί μια υπαρκτή λέξη ή μια «φτιαγμένη» (made – up) - μη λέξη (Baayen, Feldman & Schreuder, 2006). Οι χαμηλής συχνότητας λέξεις αποτελούν λέξεις οι οποίες χρειάζονται περισσότερο χρόνο και είναι δυσκολότερες στο να ανακληθούν (Yonelinas, 2002). Οι λέξεις αυτές, όμως, όταν πρόκειται για μια δοκιμασία ανάκλησης λέξεων, οδηγούν σε καλύτερα αποτελέσματα (Yonelinas, 2002).

1.3 Μέτρηση συχνότητας

Εξαιτίας του γεγονότος ότι οι μετρήσεις της συχνότητας εξαρτώνται από το μέγεθος του «σώματος» (corpus), οι ερευνητές επιλέγουν να δουλεύουν με μια σταθμισμένη μονάδα μέτρησης ώστε η ποικιλία των μετρήσεων που έχουν να μπορεί να συγκριθεί (Brysbaert et al., 2018). Μέχρι στιγμής, η κύρια σταθμισμένη μεταβλητή, η οποία χρησιμοποιείται, είναι η συχνότητα ανά εκατομμύρια λέξεις (frequency per million words - fpm) (Brysbaert et al., 2018). Οι λέξεις χαμηλής συχνότητας έχουν λιγότερα από 5 fpm ενώ οι λέξεις υψηλής συχνότητας έχουν περισσότερα από 100 fpm (Brysbaert et al., 2018). Η κλίμακα fpm παρέχει στους χρήστες λανθασμένες εντυπώσεις αναφορικά με τη συχνότητα (Brysbaert et al., 2018). Εξαιτίας αυτού, οι van Heuven, Mandera, Keuleers και Brysbaert (2014), εισηγήθηκαν την κλίμακα Zipf ως μια εναλλακτική. Αυτή η κλίμακα είναι λογαριθμική (όπως και η κλίμακα decibel) και υπολογίζεται ως $Zipf = \log_{10}$ (συχνότητα ανά εκατομμύρια λέξεις). Στην πράξη, η συγκεκριμένη κλίμακα παρέχει μετρήσεις από 1 (1 ανά 100 εκατομμύρια λέξεις) μέχρι 6 (1,000 ανά εκατομμύρια λέξεις) (van Heuven et al., 2014). Οι χαμηλής συχνότητας λέξεις αναπαρίστανται στο κάτω μισό της κλίμακας, με αριθμούς 1 μέχρι 3 ενώ οι υψηλής συχνότητας λέξεις είναι στο πάνω μισό με αριθμούς 4 μέχρι 6. Πολύ λίγες λέξεις σε μια γλώσσα έχουν υψηλότερο από 6 Zipf και συνήθως είναι οι λειτουργικές λέξεις (π.χ.: «εσύ», «με», «το») (Brysbaert et al., 2018).

1.4 Σύγκριση λέξεων υψηλής συχνότητας έναντι λέξεων χαμηλής συχνότητας

Υπάρχουν πολλές έρευνες στις οποίες γίνεται σύγκριση μεταξύ των λέξεων υψηλής συχνότητας και των λέξεων χαμηλής συχνότητας. Οι υψηλής συχνότητας λέξεις παράγονται και γίνονται γρηγορότερα αντιληπτές από τις λέξεις χαμηλής συχνότητας (Balota & Chumbley, 1984; Jescheniak & Levelt, 1994; Johnston & Barry,

2006). Σε δοκιμασίες ανάκλησης λέξεων, οι οποίες αφορούν διεργασίες της μνήμης, φαίνεται πως οι συμμετέχοντες έχουν καλύτερα αποτελέσματα με τις λέξεις υψηλής συχνότητας παρά με τις λέξεις χαμηλής συχνότητας (Balota et al., 2002). Παρά το γεγονός ότι οι λέξεις υψηλής συχνότητας ανακαλούνται ευκολότερα, είναι πιο δύσκολο να αναγνωριστούν σε δοκιμασίες που σχετίζονται με την επεισοδιακή μνήμη (Glanzer & Bowles, 1976; Yonelinas, 2002). Σύμφωνα με τους Johnston και Barry (2006) οι λέξεις υψηλής συχνότητας τυπικά κατακτούνται νωρίτερα, σε μικρότερη ηλικία, από οι χαμηλής συχνότητας λέξεις. Αυτό έχει ως αποτέλεσμα να υπάρχει μια αρνητική συσχέτιση μεταξύ της ηλικίας στην οποία αποκτούνται οι λέξεις και στη συχνότητα των λέξεων (Johnston & Barry, 2006). Ακόμη, παρατηρείται πως λέξεις οι οποίες έχουν γρηγορότερη ανταπόκριση και γρηγορότερο χρόνο αντίδρασης είναι οι λέξεις υψηλής συχνότητας οι οποίες έχουν κατακτηθεί νωρίτερα από ότι οι λέξεις χαμηλής συχνότητας οι οποίες αποκτούνται μεταγενέστερα (Pérez, 2007).

1.5 Βάσεις δεδομένων αναφορικά με τη συχνότητα εμφάνισης λέξεων

Τα τελευταία χρόνια έχει κατασκευαστεί πληθώρα ηλεκτρονικών βάσεων δεδομένων, οι οποίες παρέχουν τη συχνότητα εμφάνισης των λέξεων για πολλές γλώσσες στην Ευρώπη και τις ΗΠΑ (Kučera & Francis, 1967). Αυτές οι βάσεις δεδομένων έχουν δημιουργηθεί από το διαδίκτυο [π.χ. HAL (Lund & Burgess, 1967)], από το Google (Terzopoulos, et al., 2017). ή από τηλεοπτικούς υποτίτλους και υποτίτλους ταινιών [π.χ. SUBTLEX (New, Brysbaert, Veronis, & Pallier, 2007)] (Terzopoulos, et al, 2017). Έχει φανεί πως τα πρότυπα της συχνότητας λέξεων που βασίζονται σε υπότιτλους ταινιών, είναι πιο αποτελεσματικοί από ότι όταν είναι βασισμένες σε εφημερίδες και βιβλία (Brysbaert et al., 2018). Το γεγονός αυτό οφείλεται στο ότι, όταν οι νόρμες της συχνότητα των λέξεων στηρίζονται σε υποτίτλους ταινιών είναι πιο αντιπροσωπευτικές για τη χρήση τους στην γλώσσα που χρησιμοποιείται επί καθημερινής βάσεως (Brysbaert et al., nd). Έχουν δημιουργηθεί τέτοιες βάσεις δεδομένων και είναι διαθέσιμες για πολλές γλώσσες ανά τον κόσμο. Ένα κλασσικό παράδειγμα είναι τα Αμερικάνικα Αγγλικά που έχει ως βάση δεδομένων, για τη συχνότητα λέξεων, το SUBTLEX-US, το οποίο περιέχει 51 εκατομμύρια λέξεις, οι οποίες προέρχονται από 8,388 διαφορετικούς υποτίτλους από ταινίες και τηλεοπτικά προγράμματα (Brysbaert, New, & Keuleers, 2012). Άλλα παραδείγματα γλωσσών που

διαθέτουν παρόμοιες ηλεκτρονικές βάσεις δεδομένων είναι τα Ολλανδικά – SUBTLEX-NL (Keuleers & Brysbaert & New, 2010), τα Ελληνικά – SUBTLEX-GR (Dimitropoulou, et al., 2010), τα Κινέζικα – SUBTLEX-CH (Cai & Brysbaert, 2010), τα Ισπανικά – SUBTLEX-ESP (Cuetos, Glez-Nosti, Barbón, & Brysbaert, 2012), τα Πολωνικά – SUBTLEX-PL (Mandera, Keuleers, Wodniecka, & Brysbaert, 2015), τα Αλβανικά – SUBTLEX-AL (Avdyli, & Cuetos, 2013), τα Πορτογαλικά – SUBTLEX-PT (Soares, et al., 2015).

1.6 Βάσεις δεδομένων συχνότητας στην Ελληνική γλώσσα

Σύμφωνα με τους Terzopoulos et al. (2017), αυτή τη στιγμή, υπάρχουν τρεις βάσεις δεδομένων για την Ελληνική γλώσσα οι οποίες είναι προσβάσιμες και μπορούν να χρησιμοποιηθούν. Οι συγκεκριμένες ηλεκτρονικές βάσεις, παρέχουν τη συχνότητα των λέξεων στα Ελληνικά (Terzopoulos et al., 2017). Αρχικά, η πρώτη ευρέως διαθέσιμη βάση δεδομένων της νεοελληνικής γλώσσας, είναι το Hellenic National Corpus (HNC) (Terzopoulos et al., 2017). Έχει δημιουργηθεί από το Institute of Language and Speech Processing της Αθήνας (Terzopoulos et al., 2017). Αποτελείται από δεδομένα συλλεγόμενα από γραπτά κείμενα, στα Ελληνικά, τα οποία περιέχουν περισσότερες από 24 εκατομμύρια λέξεις (Mikros, Hatzigeorgiu & Carayannis, 2005). Τα κείμενα αυτά έχουν μαζευτεί από βιβλία, περιοδικά, εφημερίδες και άλλα έντυπα είδη τα οποία δημοσιεύτηκαν από το 1976 και μετά (Terzopoulos et al., 2017). Μέσω της ιστοσελίδας αυτής, ο χρήστης έχει τη δυνατότητα να δει λίστες με τις πιο κοινώς εμφανιζόμενες λέξεις στα Ελληνικά που περιλαμβάνονται στο HNC ή να αναζητήσει τη συχνότητα εμφάνισης μιας συγκεκριμένης λέξης (Hatzigeorgiu et al., 2000). Η δεύτερη βάση δεδομένων για την Ελληνική γλώσσα είναι το GreekLex (Kyparissiadis, van Heuven, Pitchford & Ledgeway, 2017). Το GreekLex βασίζεται κυρίως στην ορθογραφία της Νεοελληνικής γλώσσας (Ktori, van Heuven, & Pitchford, 2008). Αποτελείται από 35.304 διαφορετικούς τύπους λέξης και νεοελληνικά λήμματα, τα οποία περιλαμβάνουν τις συντακτικές και σημασιολογικές ιδιότητες καθώς και τη φωνολογική δομή κάθε λέξης (Gahl, S 2008). Από τα λήμματα αυτά, παρέχονται πληροφορίες σχετικά με το μήκος κάθε λέξης, τη συχνότητα και τη γραπτή μορφή των λέξεων αλλά και για την ορθογραφία τους (Ktori, van Heuven, & Pitchford, 2008). Η τρίτη βάση δεδομένων που αφορά την Ελληνική γλώσσα, όπως προαναφέρθηκε, είναι

το SUBTLEX-GR (Dimitropoulou, et al.,2010). Αποτελεί μια βάση δεδομένων στηριζόμενη σε ένα σύνολο 6,032 υποτίτλων στα Ελληνικά από 5,508 ταινίες και τηλεοπτικές σειρές (Dimitropoulou, et al.,2010). Έχει συνταχθεί με περισσότερες από 27 εκατομμύρια λέξεις της Νέας Ελληνικής γλώσσας. (Dimitropoulou, et al., 2010).

1.7 Χρησιμότητα συχνότητας

1.7.1 Μη κλινικοί σκοποί

Οι προαναφερθείσες βάσεις δεδομένων, που αφορούν τη συχνότητα των λέξεων, είναι χρήσιμες τόσο για μη κλινικούς όσο και για κλινικούς σκοπούς. Κάποιοι από τους σκοπούς αυτούς μπορούν να αποκαλυφθούν μέσα από τους στόχους της πιο πρόσφατης βάσης δεδομένων, που αφορά τη συχνότητα των λέξεων, στην Ελληνική γλώσσα, το HelexKids (Terzopoulos et al., 2017). Το HelexKids είναι μια ιστοσελίδα για Ελληνόφωνα παιδιά στη δημοτική εκπαίδευση (Terzopoulos et al., 2017). Έχει δημιουργηθεί με στόχο να βοηθήσει τους ερευνητές στην προσπάθεια τους να μελετήσουν την ανάπτυξη και την εξέλιξη της Ελληνικής γλώσσας (Terzopoulos et al., 2017). Η συγκεκριμένη βάση δεδομένων είναι οργανωμένη βάσει τάξεων (Α – Στ Δημοτικού), παρέχοντας έτσι τη συχνότητα στην οποία εμφανίζονται οι λέξεις ανά ηλικία και τάξη στο σχολείο. Οι λέξεις που υπάρχουν σε αυτή τη βάση δεδομένων, έχουν παρθεί από 116 βιβλία τα οποία χρησιμοποιούνται τόσο στην Ελληνική όσο και στην Κυπριακή εκπαίδευση. Μέσα από αυτά τα βιβλία, μαζεύτηκαν 68,692 ξεχωριστοί τύποι μη – λημματικών λέξεων επιλεγμένοι ανάμεσα σε 1,355, 265 στοιχεία. Το HelexKids στοχεύει στο να βοηθήσει τους εκπαιδευτικούς στην επιλογή τους για τις κατάλληλες έννοιες και λέξεις που πρέπει να διδάξουν στα παιδιά ανάλογα με την ηλικία τους. Επίσης, αποσκοπεί στο να διευκολύνει τους ερευνητές στην προσπάθεια τους να μελετήσουν την ανάπτυξη και την εξέλιξη της γλώσσας και παρέχετε ως εργαλείο χρήσης για τους συγγραφείς των εκπαιδευτικών βιβλίων για παιδιά στην Ελλάδα αλλά και στην Κύπρο (Terzopoulos et al., 2017).

Επομένως, ένας μη κλινικός σκοπός για τον οποίο μπορεί να φανεί χρήσιμο το WFE, και το να γνωρίζουμε τη συχνότητα των λέξεων μέσα από βάσεις δεδομένων στις οποίες βρίσκονται μαζεμένες οι λέξεις υψηλής και χαμηλής συχνότητας, είναι η εκμάθηση μιας γλώσσας (Hayakawa, Ning & Marian, 2020). Ακόμη ένας μη κλινικός σκοπός, που διαφαίνεται μέσω της αναφοράς στο HelexKids, είναι το ενδιαφέρον των

ερευνητών και κυρίως των ψυχολinguιστολόγων. Αυτό προκύπτει από το γεγονός ότι το HelexKids είναι η πρώτη ψυχολinguιστολογική βάση δεδομένων, που δημιουργήθηκε, και από το ότι προσφέρει στους ερευνητές την δυνατότητα να μελετούν την Ελληνική γλώσσα (Terzopoulos et al., 2017). Οι ψυχολinguιστολόγοι δείχνουν ενδιαφέρον για τις επιπτώσεις που έχουν οι γλωσσικές μεταβλητές, στις γλωσσικές συμπεριφορές μιας ομάδας χρηστών της γλώσσας, και όπως φαίνεται η συχνότητα στην οποία εμφανίζονται οι λέξεις είναι μία από αυτές (Terzopoulos et al., 2017). Ως εκ τούτου, η συχνότητα των λέξεων όπως αυτή αναπαρίσταται στον εγκέφαλο ενός ομιλητή ή ενός ακροατή φαίνεται πως έχει ιδιαίτερη σημαντικότητα. (Jescheniak & Levelt 1994) Επιπρόσθετα, η δυνατότητα που παρέχουν οι βάσεις δεδομένων που βασίζονται σε μια ποικιλία βιβλίων τα οποία διαβάζουν και χρησιμοποιούν τα παιδιά στο σχολείο επιτρέπει στους ερευνητές να έχουν μια πιο ολοκληρωμένη εικόνα όταν εξετάζουν τα φαινόμενα λεξικής επεξεργασία στα παιδιά. (Terzopoulos et al., 2017) Τέλος, οι λίστες στις οποίες φαίνεται η συχνότητα κάθε λέξης μπορούν να φανούν χρήσιμες σε μεταφραστές, διερμηνείς και σε άτομα που διδάσκουν μια γλώσσα σε ξένους μαθητές (Γούτσος, n.d.). Αποτελεί χρήσιμο εργαλείο το γεγονός πως εάν ειδικότητες σαν αυτές γνωρίζουν τη συχνότητα των λέξεων, τότε μπορούν να ξέρουν ποιες λέξεις ενδείκνυται να χρησιμοποιούνται περισσότερο από ξένους μαθητές που προσπαθούν να μάθουν την Ελληνική γλώσσα. Κατ' αυτό τον τρόπο μπορούν ευκολότερα να εστιάσουν στις βασικότερες λέξεις και έννοιες που χρειάζεται να γνωρίζει ένα άτομα ώστε να μπορεί να επικοινωνήσει αποτελεσματικά στην καθημερινότητα του. (Γούτσος, n.d.)

1.7.2 Κλινικοί σκοποί

Όσο αφορά τους κλινικούς σκοπούς στους οποίους μπορεί να φανεί χρήσιμο το να γνωρίζουμε τη συχνότητα στην οποία εμφανίζονται οι λέξεις, είναι κυρίως η αποκατάσταση κλινικών πληθυσμών (Badham, Whitney, Sanghera & Maylor, 2017). Με τη χρήση των λέξεων που συναντούνται συχνότερα σε μια γλώσσα, μπορούν να ωφεληθούν κλινικοί πληθυσμοί, όπως άτομα με αφασία (Hoffman, Rogers & Lambon Ralph, 2011). Άτομα με αφασία είναι πιθανόν να παρουσιάσουν προβλήματα και δυσκολίες σε πολλούς τομείς της γλώσσας στους οποίους θα βοηθούσε το να γνωρίζει ο ειδικός τη συχνότητα των λέξεων ώστε να εστιάσει στην αποκατάσταση τους (Nozari, Kittredge, Dell & Schwartz, 2010). Κάποιοι από αυτούς τους τομείς είναι η σημασιολογία και η παραγωγή σημασιολογικών εννοιών, η κατονομασία εικόνων,

καθώς μπορεί να εμφανίσουν και φωνολογικά λάθη (Howes, 1964). Όπως αναφέρεται από τους Kittredge et al. (2008), μέσα από την έρευνα τους σε άτομα με αφασία, βρέθηκε πως οι λέξεις με υψηλότερη συχνότητα παρουσίασαν λιγότερα σημασιολογικά και φωνολογικά λάθη. Ωστόσο, φαίνεται πως η επίδραση που έχει η συχνότητα στα σημασιολογικά λάθη, έχει αποδειχθεί αρκετά μικρότερη αλλά και λιγότερο ισχυρή από ότι στα φωνολογικά σφάλματα. (Nozari, Kittredge, Dell, & Schwartz, 2010)

Παράλληλα άλλοι σκοποί των λέξεων υψηλής και χαμηλής συχνότητας, είναι πως θα μπορούσαν να επωφεληθούν κλινικοί πληθυσμοί όπως άτομα με εγκεφαλικό. Οι Hoffman et al., (2011) υποστηρίζουν πως άτομα που υπέστηκαν εγκεφαλικό παρουσιάζουν προβλήματα σχετικά με τη συχνότητα των λέξεων. Επίσης, αναφέρουν πως άτομα με σημασιολογική άνοια λόγω του γεγονότος ότι μετά από ατροφία του πρόσθιου κροταφικού λοβού που υπέστηκαν, η εννοιολογική τους γνώση έχει υποβιβαστεί με αποτέλεσμα να παρουσιάζουν έντονα φαινόμενα στη συχνότητα των λέξεων. Ως εκ τούτου, εάν οι κλινικοί γνωρίζουν τη συχνότητα των λέξεων, τόσο τις λέξεις υψηλής όσο και τις λέξεις χαμηλής συχνότητας, μπορούν να γνωρίζουν που ακριβώς να εστιάσουν ώστε να βοηθήσουν τους ασθενείς τους να παρουσιάσουν πρόοδο (Hoffman, Rogers & Lambon Ralph, 2011). Ένας ακόμα τομέας που θα μπορούσε να αποκτήσει πολλά πλεονεκτήματα από τις λέξεις υψηλής και χαμηλής συχνότητας, είναι τα άτομα που παρουσιάζουν διαταραχές στη μνήμη, δηλαδή άτομα που πάσχουν από τη νόσο Αλτσχάιμερ ή από μια Ήπια Γνωστική Διαταραχή (Nikolaev et al., 2018). Τα άτομα αυτά συχνά δυσκολεύονται να ανακτήσουν λέξεις και έχουν προβλήματα στις λεξικές αποφάσεις (Balota, Burgess, Cortese & Adams, 2002). Έτσι, οι κλινικοί που ασχολούνται με τέτοιου είδους περιστατικά μπορούν να εστιάσουν στο να εκπαιδεύσουν αυτά τα άτομα με τις λέξεις υψηλής συχνότητας, ούτως ώστε να μπορούν να επικοινωνούν με τους γύρω τους (Nikolaev et al., 2018).

Επιπρόσθετα, υπάρχουν και άλλες διαταραχές στις οποίες μπορεί να φανεί χρήσιμο οι κλινικοί ή οι ειδικοί να γνωρίζουν τις λέξεις υψηλής και χαμηλής συχνότητας προκειμένου να τις εντάξουν στα προγράμματα αποκατάστασης τους. Όπως αναφέρεται από τον Howes (1964), μια διαταραχή στην οποία μπορούν να χρησιμοποιηθούν οι λέξεις χαμηλής και υψηλής συχνότητας ως εργαλείο αποκατάστασης, είναι στον τραυλισμό. Ο Howes (1964) εκφράζει πως οι λέξεις στις οποίες παρουσιάζονται περισσότερες δυσρυθμίες είναι οι λέξεις χαμηλής συχνότητας.

Βάσει αυτού, ένας ειδικός για να παρέχει βοήθεια και τρόπους αντιμετώπισης σε ένα παιδί ή ένα ενήλικα με τραυλισμό, χρειάζεται να γνωρίζει τόσο τις λέξεις υψηλής όσο και τις λέξεις χαμηλής συχνότητας (Howes, 1964).

Επίσης, εκτός από τα άτομα με τραυλισμό, μπορούν να βοηθηθούν και τα παιδιά σε σχολική ηλικία ώστε να αποκτήσουν ευχέρεια στην ανάγνωση τους (Δημητριάδου, n.d.). Αυτό μπορεί να προκύψει αφού όταν οι ειδικοί έχουν εν γνώσει τις λέξεις υψηλής συχνότητας μπορούν να επιλέξουν καλύτερα ποιο κείμενο να δώσουν στα παιδιά ώστε να μπορούν να αναγνωρίζουν και να διαβάζουν τις λέξεις αυτές (Δημητριάδου, n.d.). Γνωρίζοντας τις λέξεις υψηλής συχνότητας, οι οποίες συναντούνται περισσότερες φορές σε ένα κείμενο ή στην καθημερινή ομιλία, οι ειδικοί μπορούν να εστιάσουν σε αυτές ώστε να βοηθήσουν τα παιδιά να μάθουν να τις χρησιμοποιούν ορθά τόσο στον προφορικό όσο και στον γραπτό λόγο, δηλαδή, να τις γράφουν ορθογραφημένα (Δημητριάδου, n.d.). Επιπλέον, όταν ένας ειδικός ξέρει τις λέξεις υψηλής συχνότητας σε μια γλώσσα, μπορεί να παρέχει βοήθεια σε παιδιά με μαθησιακές δυσκολίες και δυσλεξία (Δημητριάδου, n.d.). Παιδιά με τέτοιες διαταραχές μπορούν να εκπαιδευτούν για να βελτιώσουν την αυτόματη αναγνώριση των λέξεων και να αυξήσουν την ταχύτητα ανάγνωσης τους. Ακόμη, μέσω των λέξεων υψηλής συχνότητας, ένας ειδικός μπορεί να βοηθήσει τα παιδιά ώστε να εμπλουτίσουν το λεξιλόγιο τους. (Δημητριάδου, n.d.)

Σύμφωνα με τους Jiménez González και Hernández Valle (2000), οι οποίοι διεξήγαν μια έρευνα όπου έκαναν μια σύγκριση της ικανότητας κατονομασίας των παιδιών με μαθησιακές δυσκολίες, τα οποία είχαν διαγνωσθεί με δυσκολίες στην εύρεση λέξεων, με τυπικά αναπτυσσόμενα παιδιά και άτομα που επίσης παρουσίαζαν μαθησιακές δυσκολίες αλλά μετά από αξιολόγηση τους, φάνηκε πως έχουν καλές ικανότητες στην έκφραση της γλώσσας. Ακολούθως, μέσω των λέξεων υψηλής και χαμηλής συχνότητας υπολόγισαν τις δεξιότητες κατονομασίας των δύο ομάδων, τις ικανότητες εντοπισμού των λέξεων – στόχων βάσει της συχνότητας εμφάνισης των λέξεων που χρησιμοποιήθηκαν καθώς και τον χρόνο αντίδρασης σε αυτές τις διαδικασίες (Jiménez González & Hernández Valle, 2000). Το αποτέλεσμα που βρέθηκε μέσα από τη σύγκριση των ομάδων που δημιουργήθηκαν, είναι το ότι η ομάδα που παρουσίαζε μαθησιακές δυσκολίες και προβλήματα στην εύρεση λέξεων εκδήλωσε περισσότερες δυσκολίες σε δραστηριότητες που αφορούσαν τον χρόνο αντίδρασης, την

κατονομασία και τον εντοπισμό λέξεων από ότι η δεύτερη ομάδα (Jiménez González & Hernández Valle, 2000). Ακόμη, τα παιδιά με μαθησιακές δυσκολίες και προβλήματα στην εύρεση των λέξεων, παρουσίασαν περισσότερα σφάλματα στις λέξεις χαμηλής συχνότητας και στις τρεις συνθήκες (Newman German, 1979). Επομένως, μέσω των ευρημάτων της έρευνας αυτής φαίνεται πως χρειάζεται οι ειδικοί να γνωρίζουν τις λέξεις χαμηλής και υψηλής συχνότητας έτσι ώστε να μπορέσουν να βοηθήσουν τα παιδιά με μαθησιακές δυσκολίες να βελτιώσουν τις ακαδημαϊκές τους επιδόσεις (Jiménez González & Hernández Valle, 2000). Όπως σημειώνεται στην συγκεκριμένη έρευνα, η συχνότητα εμφάνισης των λέξεων είναι μια σημαντική μεταβλητή όσο αφορά τον χρόνο αντίδρασης και την εύρεση λέξεων (Perez, 2007). Οπότε, εξάγεται το συμπέρασμα πως γνωρίζοντας τις λέξεις υψηλής και χαμηλής συχνότητας τα παιδιά με μαθησιακές δυσκολίες μπορούν εκπαιδευτούν ώστε να χρειάζονται λιγότερο χρόνο αντίδρασης και επομένως να έχουν καλύτερες επιδόσεις στο σχολείο (Jiménez González & Hernández Valle, 2000). Όπως υποστηρίζουν οι Nation, Marshall, & Snowling (2001), όταν οι κλινικοί έχουν εις γνώσιν την συχνότητα στην οποία εμφανίζονται οι λέξεις, τότε μπορούν να βοηθηθούν και παιδιά με δυσλεξία. Παιδιά και κυρίως μαθητές που έχουν διαγνωσθεί με δυσλεξία παρουσιάζουν πολλές φωνολογικές αλλά και σημασιολογικές δυσκολίες και αυτό παρατηρείται κυρίως στις λέξεις χαμηλής συχνότητας (Nation, Marshall, & Snowling, 2001).

Τέλος, ένας ακόμη λόγος για τον οποίο είναι σημαντικό οι ειδικοί να γνωρίζουν τη συχνότητα στην οποία εμφανίζονται οι λέξεις, είναι το να βοηθήσουν άτομα που δεν έχουν λεκτική επικοινωνία να μπορέσουν να εκφράζονται με τους γύρω τους (LIU & SLOANE, 2006). Αυτό μπορεί να επιτευχθεί μέσω της εναλλακτικής και επαυξητικής επικοινωνίας (LIU & SLOANE, 2006). Ένας κλινικός χρειάζεται να ενσωματώσει στην συσκευή επικοινωνίας ενός ατόμου χωρίς λεκτική επικοινωνία, σημασιολογικές κατηγορίες και το βασικό λεξιλόγιο από την κάθε κατηγορία (Hill & Romich, 2000). Όταν πρόκειται για μια συσκευή επικοινωνίας, το βασικό λεξιλόγιο αναφέρεται στις λέξεις Β. Α. (που είναι απαραίτητες για το άτομο που θα τη χρησιμοποιεί (Hill & Romich, 2000). Οι κατηγορίες αυτές πρέπει να αντανakλούν την καθημερινή επικοινωνία του ατόμου, είτε είναι ενήλικας είτε είναι παιδί (LIU & SLOANE, 2006). Υπάρχουν δύο σημαντικοί λόγοι για τους οποίους το βασικό λεξιλόγιο αποτελεί ένα αναπόσπαστο κομμάτι μιας συσκευής επικοινωνίας (LIU & SLOANE, 2006). Πρώτον,

επιτρέπει στα άτομα που το χρησιμοποιούν να παράγουν το μεγαλύτερο ποσοστό των λέξεων που απαιτούνται για κάθε έκφραση η οποία συναντάται επί καθημερινής βάσεως, όπως αποκαλούνται και αλλιώς, το λειτουργικό λεξιλόγιο (LIU & SLOANE, 2006). Δεύτερον, μέσω της ύπαρξης του βασικού λεξιλογίου στη συσκευή επικοινωνίας, με τη βοήθεια ενός συνόλου ουσιαστικών, το άτομο σε περίπτωση που οι λέξεις που χρειάζεται ώστε να εκφράσει μια ανάγκη του λείπουν από το σύστημα εναλλακτικής και επαυξητικής επικοινωνίας που χρησιμοποιεί, τότε μπορεί να χρησιμοποιήσει τα βασικά ουσιαστικά ώστε να τα αντικαταστήσει (LIU & SLOANE, 2006). Συνεπώς, ο κλινικός χρειάζεται να ξέρει τις λέξεις που συναντούνται συχνότερα στην γλώσσα του ατόμου που θέλει να χρησιμοποιεί το σύστημα εναλλακτικής και επαυξητικής τεχνολογίας ώστε να του προσφέρει όσο το δυνατό πιο αποτελεσματικότερη επικοινωνία. (LIU & SLOANE, 2006)

1.8 Διαφορές μεταξύ της Νέας Ελληνικής γλώσσας και της Κυπριακής διαλέκτου

Παρά το γεγονός ότι, όπως προαναφέρθηκε, το HelexKids αποτελεί μια βάση δεδομένων η οποία αφορά τα παιδιά στη Δημοτική Εκπαίδευση τόσο στην Ελλάδα όσο και στην Κύπρο, οι λέξεις που αποτελούν αυτή την ιστοσελίδα είναι στα Ελληνικά. Πολλές από τις λέξεις που υπάρχουν στην συγκεκριμένη ιστοσελίδα, χρησιμοποιούνται από τα παιδιά στην Κύπρο (Terzopoulos, et al., 2017). Παρ' όλα αυτά υπάρχουν αρκετές διαφορές ανάμεσα στην Νέα Ελληνική γλώσσα και την Κυπριακή διάλεκτο (Pavlou & Paparavliou, 2004). Όπως αναφέρεται από τους Sophocleous και Wilks (2010), η Νέα Ελληνική γλώσσα και η Κυπριακή διάλεκτος διαφέρουν φωνολογικά, μορφολογικά, λεξικά και συντακτικά. Παρατηρούνται διαφορές στην φωνολογία, αφού στην Κυπριακή διάλεκτο κάποια από τα σύμφωνα (π.χ.: /k/ και /x/) αλλάζουν ή διπλασιάζονται και μερικά φωνήματα υφίστανται αλλοιώσεις (Pavlou & Paparavliou, 2004). Ακόμη, υπάρχουν διαφορές στην μορφολογία αφού στην Κυπριακή διάλεκτο χρησιμοποιείται το πρόθεμα –ε ώστε να δηλωθεί παρελθοντικός χρόνος, για τον πληθυντικό αριθμό τρίτου ατόμου χρησιμοποιείται το / usin / ως κατάληξη, ενώ παράλληλα προστίθεται τελικό –ν σε πολλές λέξεις (Pavlou & Paparavliou, 2004). Διαφορές παρουσιάζονται και στην σύνταξη, κυρίως στη θέση των κλιτικών αλλά και στη σημασιολογία ή το λεξιλόγιο αφού ένας μεγάλος αριθμός των λέξεων της

Κυπριακής διαλέκτου προέρχονται από τα Τούρκικα, τα Αραβικά, τα Αγγλικά, τα Ιταλικά και τα Γαλλικά (Ραυλού & Ραπαρανίου, 2004).

1.9 Σκοπός της παρούσας έρευνας

Βάσει όλων αυτών των διαφορών που προαναφέρθηκαν ανάμεσα στην Ελληνική γλώσσα και την Κυπριακή διάλεκτο, οι ήδη υπάρχουσες βάσεις δεδομένων στα Ελληνικά, δεν μπορούν να καλύψουν πλήρως τις ανάγκες τόσο των παιδιών ή των ενηλίκων όσο και των ειδικών υγείας στην Κύπρο. Επομένως, φαίνεται πως υπάρχει ένα κενό όσο αφορά αυτό τον τομέα στην Κυπριακή διάλεκτο. Όπως αναφέρθηκε και προηγουμένως, υπάρχουν πολλά πλεονεκτήματα, τόσο κλινικά όσο και μη κλινικά, από την ύπαρξη μιας βάσης δεδομένων η οποία παρέχει την συχνότητα εμφάνισης των λέξεων σε μια γλώσσα. Έτσι, στόχος της εν λόγω έρευνας είναι η καταμέτρηση συχνότητας λέξεων στην Κυπριακή διάλεκτο με σκοπό τη δημιουργία της πρώτης λίστας συχνότητας λέξεων σημασιολογικού περιεχομένου στον Κυπριακό πληθυσμό. Με τη δημιουργία μιας τέτοιας λίστας μπορούν να επωφεληθούν τόσο οι ενήλικες όσο και τα παιδιά. Τέλος, είναι σημαντικό να σημειωθεί πως η δημιουργία αυτής της λίστας δεδομένων θα αποτελέσει ένα σημαντικό εργαλείο στη δημιουργία γνωστικών εργαλείων αξιολόγησης και αποκατάστασης κλινικών πληθυσμών.

2 Μεθοδολογία Έρευνας

2.1 Δείγμα και κριτήρια εισαγωγής στην έρευνα

Όπως προαναφέρθηκε, ο σκοπός της εν λόγω έρευνας είναι η καταμέτρηση συχνότητας σημασιολογικών λέξεων στον Κυπριακό πληθυσμό, με στόχο τη δημιουργία μιας βάσης δεδομένων η οποία θα στηρίζεται στη συχνότητα εμφάνισης των λέξεων. Η παρούσα έρευνα πραγματοποιήθηκε με συνολικό δείγμα 10 ατόμων. Έγινε σκόπιμη δειγματοληψία αφού τα επιλεγόμενα άτομα έπρεπε να πληρούν κάποια κριτήρια ώστε να αποτελούν αντιπροσωπευτικό δείγμα. Στην έρευνα συμμετείχαν άτομα ηλικίας εξήντα ετών και άνω τα οποία έπρεπε να έχουν την Κυπριακή ως μητρική τους γλώσσα. Επίσης, τα άτομα αυτά έπρεπε να έχουν σημειώσει συνολική βαθμολογία τουλάχιστον 27 στο διαγνωστικό εργαλείο Mini – Mental State Examination (MMSE). Άτομα τα οποία δεν πληρούσαν τα πιο πάνω κριτήρια δεν συμπεριλήφθηκαν στην έρευνα.

2.2 Διαδικασία συλλογής δεδομένων

Η συλλογή δεδομένων πραγματοποιήθηκε στο σπίτι του κάθε συμμετέχοντα, με προσωπικές συνεντεύξεις. Αρχικά, στα επιλεγόμενα άτομα παρουσιαζόταν λεκτικά η διαδικασία που θα ακολουθείτο. Ακολούθως, τους δινόταν ένα έντυπο ενημέρωσης όπου αναφέρονταν όλες οι πληροφορίες σχετικά με την έρευνα (Σκοπός, Φύση της συμμετοχής, Κίνδυνοι συμμετοχής, Εμπιστευτικότητα, Διαδικασία συμμετοχής στην έρευνα, Δικαίωμα μη συμμετοχής ή απόσυρσης) και ένα έντυπο συγκατάθεσης για τη συμμετοχή τους στην έρευνα. Εφόσον τα άτομα ήταν σύμφωνα στο να αποτελούν μέρος της παρούσας έρευνας, υπέγραφαν και τα δύο έντυπα.

Στη συνέχεια, οι συμμετέχοντες κλήθηκαν να συμπληρώσουν ένα έντυπο δημογραφικών στοιχείων. Στο έντυπο αυτό οι συμμετέχοντες έπρεπε να δηλώσουν το φύλο, την ηλικία, το μορφωτικό τους επίπεδο, το επάγγελμα, τον τόπο διαμονής τους, την υπηκοότητα και εάν γνωρίζουν κάποια άλλη γλώσσα.

2.3 Μέσα και τρόπος συλλογής δεδομένων

2.3.1 Mini Mental State Examination

Στους συμμετέχοντες χορηγήθηκε το Mini Mental State Examination (MMSE). Το MMSE αποτελεί ένα ευρέως χρησιμοποιούμενο εργαλείο το οποίο ανιχνεύει την έκπτωση ή την εξασθένηση των γνωστικών λειτουργιών σε ενήλικες (Loewenstein, et al., 2000). Είναι ένα απλό και σύντομο εργαλείο, καθώς διαρκεί από 5 μέχρι 10 λεπτά (Μπουρονίκου, Παπαντωνίου, Μωραΐτου & Σαρρής, 2014). Αποτελείται από 30 σημεία τα οποία χωρίζονται σε 6 υποκατηγορίες οι οποίες και εξετάζονται (Loewenstein, et al., 2000). Οι υποκατηγορίες αυτές, περιλαμβάνουν ερωτήσεις αναφορικά με τον προσανατολισμό, την προσοχή, την μνήμη και την ανάκληση και την γλώσσα (Galea & Woodward, 2005). Αρχικά εξετάζεται ο προσανατολισμός στον χρόνο και τον χώρο μέσω ερωτήσεων για την ημερομηνία και τον χώρο στον οποίο βρίσκεται το άτομο (Loewenstein, et al., 2000). Στη συνέχεια, εξετάζεται η βραχύχρονη μνήμη αφού ο εξεταστής ονομάζει τρία αντικείμενα και το άτομο καλείται να τα επαναλάβει αλλά και να τα θυμάται για τη συνέχεια (Loewenstein, et al., 2000). Ακολούθως, αξιολογείται η προσοχή και η ικανότητα υπολογισμών, αφού το άτομο καλείται να κάνει αφαιρέσεις και να συλλαβίσει ανάποδα μια λέξη (Μπουρονίκου, Παπαντωνίου, Μωραΐτου & Σαρρής, 2014). Μετά, εξετάζεται η ικανότητα ανάκλησης λέξεων αλλά και η γλώσσα μέσω ονομασίας αντικειμένων, επανάληψης προτάσεων, ακολουθίας εντολών, ανάγνωσης, γραφής και αντιγραφής σχήματος (Μπουρονίκου, Παπαντωνίου, Μωραΐτου & Σαρρής, 2014). Η υψηλότερη βαθμολογία που μπορεί να επιτύχει κάποιος στο MMSE είναι 30 βαθμοί (Μπουρονίκου, Παπαντωνίου, Μωραΐτου & Σαρρής, 2014). Βαθμολογία ίση ή μικρότερη από 23 βαθμούς, αποτελεί μια ένδειξη γνωστικής έκπτωσης η οποία μπορεί να σχετίζεται με διαταραχές εξασθένησης των γνωστικών λειτουργιών (π.χ. άνοια) (Loewenstein, et al., 2000).

Το συγκεκριμένο διαγνωστικό εργαλείο χορηγήθηκε στους συμμετέχοντες με σκοπό να αξιολογηθεί η γνωστική τους λειτουργία. Επομένως, αυτό το διαγνωστικό εργαλείο χορηγήθηκε προκειμένου να είμαστε σίγουροι πως όλοι οι ενήλικες που θα συμπεριλαμβάνονταν στην έρευνα θα είναι γνωστικά υγιείς, ώστε να εξασφαλίσουμε τα καλύτερα δυνατά αποτελέσματα μέσω των απαντήσεων των συμμετεχόντων.

2.3.2 Κύριο ερευνητικό εργαλείο

Η συλλογή των δεδομένων έγινε με τη χρήση μιας λίστας συχνότητας λέξεων σημασιολογικού περιεχομένου (Παράρτημα Ι). Η λίστα αυτή δημιουργήθηκε από την ερευνητική ομάδα με σκοπό να αξιολογήσει τη συχνότητα εμφάνισης των λέξεων στον Κυπριακό πληθυσμό. Συγκεκριμένα, η λίστα καταγραφής της συχνότητας, αφορούσε σημασιολογικές κατηγορίες οι οποίες συναντούνται πιο συχνά στην καθημερινότητα των ατόμων στον Κυπριακό πληθυσμό. Η λίστα αυτή αποτελείτο από 12 σημασιολογικές κατηγορίες.

Στους συμμετέχοντες αρχικά δινόταν η οδηγία: «Θα σας αναφέρω μια σημασιολογική κατηγορία και θέλω να προσπαθήσετε να μου πείτε όσες περισσότερες λέξεις πιστεύετε πως ταιριάζουν στην κατηγορία αυτή.» Δεν υπήρχε κάποιος περιορισμός ως προς τον χρόνο που δινόταν στους συμμετέχοντες. Έτσι, οι συμμετέχοντες έπρεπε να ανακαλέσουν ελεύθερα όσες περισσότερες λέξεις μπορούσαν να θυμηθούν από την κύρια σημασιολογική κατηγορία που τους αναφερόταν. Αφού γινόταν η ελεύθερη ανάκληση, για κάποιες από τις σημασιολογικές ομάδες δίνονταν στους συμμετέχοντες κάποιες υποκατηγορίες με σκοπό τη διευκόλυνση της ανάκλησης. Οι κατηγορίες και υπόκατηγορίες ήταν οι εξής:

1. Ζώα: κατοικίδια – σαρκοφάγα – ψάρια – πουλιά – έντομα – φάρμας.
2. Πράγματα που αγοράζεις στην Υπεραγορά
3. Φαγητά/Πράγματα που τρώμε: Γλυκά – Μπαχαρικά – Τυριά.
4. Λαχανικά: Σαλάτας – Πράσινα – Με φύλλα.
5. Ρούχα: Γυναικεία – Αντρικά – Καλοκαιρινά – Χειμερινά.
6. Αντικείμενα/Πράγματα: Στην κουζίνα – Στο γραφείο – Έπιπλα – Εργαλεία.
7. Ποτά/Ροφήματα: Οινοπνευματώδη – Ροφήματα που πίνουμε το χειμώνα/
Ροφήματα που πίνουμε το καλοκαίρι.
8. Μουσικά Όργανα
9. Οχήματα
10. Μέρη/Κτήρια
11. Λουλούδια

12. Χώροι σπιτιού

Για τη συμπλήρωση της λίστας αυτής, οι συμμετέχοντες έλεγαν ανακαλούσαν λέξεις που αντιστοιχούσαν σε κάθε μια από τις κατηγορίες και οι απαντήσεις καταγράφονταν. Η συνολική διάρκεια χορήγησης της λίστας ήταν περίπου μια ώρα. Ακολούθως, κάθε συμπληρωμένη λίστα συχνότητας σημασιολογικών λέξεων, κάθε έντυπο καταγραφής του διαγνωστικού εργαλείου (MMSE) που χορηγήθηκε, κάθε φόρμα δημογραφικών δεδομένων καθώς και κάθε έντυπο συγκατάθεσης πήραν από ένα αριθμό συμμετέχοντα, με τον οποίο ταξινομήθηκαν. Από τα στοιχεία που συλλέχθηκαν, δεν χρησιμοποιήθηκαν τα ονόματα των συμμετεχόντων αλλά σε κάθε έναν δινόταν ένας κωδικός ώστε να μπορεί να γίνει ταυτοποίηση.

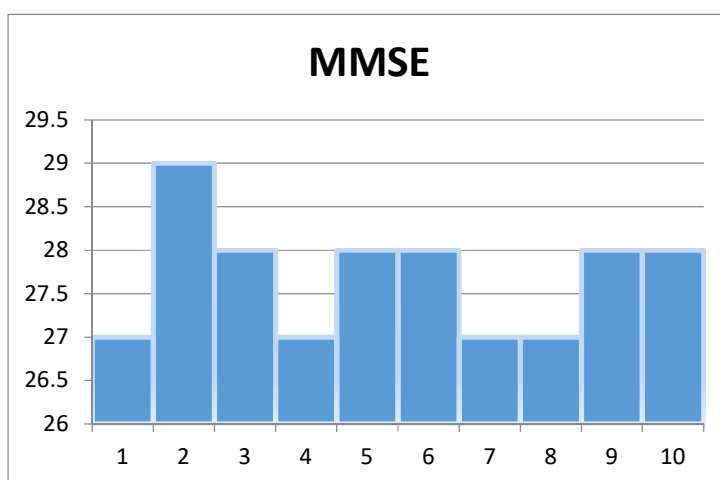
Η στατιστική ανάλυση των δεδομένων έγινε με την χρήση του στατιστικού προγράμματος Jamovi.

3 Αποτελέσματα

Όλα τα δεδομένα που συλλέχθηκαν αναλύθηκαν στο στατιστικό λογισμικό Jamovi μέσω περιγραφικής στατιστικής.

3.1 Περιγραφή δείγματος

Το δείγμα της παρούσας έρευνας αποτελείται από 10 άτομα. Από αυτά τα 10 άτομα οι 4 ήταν γυναίκες (40%) και οι 6 ήταν άνδρες (60%). Ο μέσος όρος ηλικίας των ατόμων ήταν 72,7. Ο τόπος διαμονής των συμμετεχόντων ήταν η Λεμεσός (90%) και η Πάφος (10%). Βάση της ανάλυσης του γνωστικού εργαλείου MMSE, ο μέσος όρος βαθμολογίας των συμμετεχόντων ήταν 27,7. Η μεγαλύτερη τιμή ήταν 29 και η μικρότερη 27. Συγκεκριμένα, ο μέσος όρος βαθμολογίας των γυναικών στο MMSE ήταν 27,5 ενώ ο μέσος όρος βαθμολογίας των ανδρών ήταν 28.



Διάγραμμα 1: Βαθμολογίες συμμετεχόντων στο γνωστικό εργαλείο MMSE

3.2 Περιγραφικά Ευρήματα

Από την περιγραφική στατιστική ανιχνεύθηκαν συνολικά 521 λέξεις σημασιολογικού περιεχομένου (Παράρτημα II). Από αυτές, οι 374 αφορούν λέξεις χαμηλής συχνότητας, οι 81 λέξεις μέτριας συχνότητας και οι υπόλοιπες 66 λέξεις υψηλής συχνότητας. Όλες οι λέξεις που συλλέχθηκαν μέσω της λίστας που δημιουργήσαμε, χωρίστηκαν σε τρεις κατηγορίες ανάλογα με την συχνότητα εμφάνισης

τους. Ως λέξεις χαμηλής συχνότητας καταγράφηκαν οι λέξεις που ειπώθηκαν από 1 μέχρι 3 άτομα. Ως μέτριας συχνότητας λέξεις, καταγράφηκαν όσες αναφέρθηκαν από 4 έως 6 άτομα. Ως λέξεις υψηλής συχνότητας, καταγράφηκαν οι λέξεις που ειπώθηκαν από 7 μέχρι 10 συμμετέχοντες. Με αυτό τον τρόπο βρέθηκαν οι λέξεις που χρησιμοποιούνται συχνότερα, οι λέξεις με μέτρια συχνότητα εμφάνισης, καθώς και οι λέξεις που χρησιμοποιούνται λιγότερο σε κάθε σημασιολογική κατηγορία.

3.3 Συχνότητα λέξεων ανά σημασιολογική κατηγορία

3.3.1 Υψηλής Συχνότητας Λέξεις

Για τη σημασιολογική κατηγορία Ζώα, οι δέκα πιο συχνά χρησιμοποιούμενες λέξεις είναι: σκύλος (10), γάτα (10), πρόβατο (10), λιοντάρι (10), κατσαρίδα (9), ελέφαντας (9), φίδι (9), γάιδαρος (9), κότα (8), αγελάδα (8). Στην σημασιολογική κατηγορία Υπεραγορά, οι πιο συχνά χρησιμοποιούμενες λέξεις είναι: βούτυρος (8), κρεμμύδια (8), καφές (8), γάλα (7), ζάχαρη (7), πατάτες (7). Όσο αφορά τη σημασιολογική κατηγορία Φαγητά, οι συχνότερα χρησιμοποιούμενες λέξεις είναι: σουβλά (9), φασόλες (9), κουπέπια (8), πιπέρι (8), κασκαβάλι (8), χαλούμι (7), μακαρόνια (7), ψητό (7). Στη σημασιολογική κατηγορία Λαχανικά, οι υψηλής συχνότητας λέξεις είναι: ρόκα (10), κραμπί (10), μαρούλι (9), ντομάτα (8), αγγούρι (8), σπανάκι (7), σέλινο (7), λάχανα (7). Αναφορικά με τη σημασιολογική κατηγορία Ρούχα, οι πιο συχνά χρησιμοποιούμενες λέξεις είναι: κλάτσες (10), παντελόνι (9), σακάκι (9), φουστάνι (9), φανέλα (8), φούστα (8), τρικό (7), πουκάμισο (7), ζακέτα (7). Στην τελευταία σημασιολογική κατηγορία, Αντικείμενα, οι 10 συχνότερα χρησιμοποιούμενες λέξεις είναι: τραπέζι (10), σφυρί (9), κατσαβίδι (9), καρέκλες (9), κατσαρόλα (9), μολύβι (8), κρεβάτι (8), κομπιούτερ (8), πιρούνι (8), μαχαίρι (8).

3.3.2 Μέτριας Συχνότητας Λέξεις

Δέκα από τις λέξεις που βρέθηκαν να έχουν μέτρια συχνότητα εμφάνισης στην σημασιολογική κατηγορία Ζώα, είναι: γουρούνι (6), λυθρίνι (6), καμηλοπάρδαλη (6), τίγρης (6), γάλος (5), αλεπού (5), πέστροφα (5), πίθηκος (5), σουπιά (5), μέλισσα (5). Στη σημασιολογική κατηγορία Υπεραγορά, δέκα από τις λέξεις μέτριας συχνότητας είναι: ψωμί (6), φασόλια (6), αγγούρι (6), ντομάτα (6), μακαρόνια (5), σοκολάτες (5), κρέας (5), μαρούλι (5), λούντζα (4), ρύζι (4). Σχετικά με την σημασιολογική κατηγορία

Φαγητά, δέκα από τις λέξεις που βρέθηκαν να έχουν μέτρια συχνότητα είναι: αναρή (6), δάφνη (6), σουβλάκια (6), λουβί (6), γαλακτομπούρεκο (6), τούρτα (5), φέτα (5), γλυκά του κουταλιού (5), ρίγανη (5), μακαρόνια του φούρνου (5). Όσο αφορά τη σημασιολογική κατηγορία Λαχανικά, ως μέτριας συχνότητας λέξεις καταγράφηκαν οι: μπρόκολο (6), μαϊντανός (6), κόλιανδρος (5), γλιστρίδα (5), πιπεριές (4), δυόσμος (4). Στην σημασιολογική κατηγορία Ρούχα, βρέθηκαν ως μέτριας συχνότητας λέξεις οι: παντόφλες (6), εσώρουχα (6), παπούτσια (5), μαγιό (5), παλτό (4), μπλούζα (4). Αναφορικά με τη σημασιολογική κατηγορία Αντικείμενα, δέκα από τις λέξεις που είχαν μέτρια συχνότητα εμφάνισης είναι: γραφείο (6), ψυγείο (6), βιβλίο (6), τηλεόραση (6), φάιλ (5), σμυρίλιος (5), μέτρο (4), μαξιλάρι (4), κομοδίνο (4), τσάντα (4).

3.3.3 Χαμηλής Συχνότητας Λέξεις

Βάσει των δεδομένων στη σημασιολογική κατηγορία Ζώα, 10 από τις λέξεις χαμηλής συχνότητας είναι: αετός (3), λαβράκι (3), πεζούνη (3), λίμπουρος (3), σφήκας (3), κοράζινος (3), μπαρμπούνη (3), γορίλλας (3), ελάφι (3), γόπα (3). Κάποιες από τις λέξεις χαμηλής συχνότητας στην σημασιολογική κατηγορία Υπεραγορά, καταγράφηκαν οι εξής λέξεις: σκόνη πλυσίματος (3), σαμπουάν (3), λουβί (3), σαλάμι (3), γιαούρτι (3), κάρβουνα (3), μπισκότα (3), κριθαράκι (3), κουκιά (3), μαϊντανός (2). Στη σημασιολογική κατηγορία Φαγητά, δέκα από τις λέξεις με χαμηλή συχνότητα εμφάνισης είναι: σοκολατίνα (3), ψωμί (3), δυόσμος (3), κοτόπουλο (3), τυρί (3), ρεβύθια (3), κολοκάσι (3), σεφταλιές (3), γλυκό καρυδάκι (3), πατάτες (3). Μερικές από τις λέξεις χαμηλής συχνότητας στη σημασιολογική κατηγορία Λαχανικά είναι: λάχανα (3), κρεμμύδι (3), κουνουπίδι (3), καρότο (3), ραπανάκι (3), κρεμμύδι φρέσκο (2), λουβάνα (2), κουλούμπρα (2), καπάρι (2), ραδίκια (1). Σχετικά με τη σημασιολογική κατηγορία Ρούχα, δέκα από τις λέξεις χαμηλής συχνότητας είναι: γάντια (3), καπέλο (3), σιάλπα (3), ταγιέρ (2), σκούφος (3), καπαρντίνα (2), φόρμες (2), τελαντωτή (2), πουλόβερ (1), ρόπα (1). Αναφορικά με την τελευταία σημασιολογική κατηγορία, τα Αντικείμενα, κάποιες από τις λέξεις με χαμηλή συχνότητα εμφάνισης είναι: καρφίτσες (3), ψαλίδι (3), κουβέρτα (3), σινί (3), τετράδιο (3), κούσπος (3), τσάπα (3), σιγάτσα (3), πόρτα (3), σίδερο (3).

4 Συζήτηση

Η παρούσα έρευνα πραγματοποιήθηκε με σκοπό την καταμέτρηση της συχνότητας των λέξεων που χρησιμοποιούνται στην Κυπριακή διάλεκτο. Απώτερος στόχος ήταν η δημιουργία της πρώτης λίστας συχνότητας λέξεων σημασιολογικού περιεχομένου, στον Κυπριακό πληθυσμό με την οποία θα έχουν όφελος τόσο οι ενήλικες όσο και τα παιδιά.

Στις λέξεις υψηλής συχνότητας, όπως ορίσαμε αυτές που παρουσιάστηκαν 7 έως 10 φορές, παρατηρούμε λέξεις οι οποίες είναι ευρέως γνωστές και αναφέρονται συχνότερα στον Κυπριακό πληθυσμό. Παρατηρούμε ότι η κατηγορία «Ζώα» περιλαμβάνει 19 λέξεις και η κατηγορία «Αντικείμενα» περιλαμβάνει 16 λέξεις. Εν αντιθέσει οι υπόλοιπες κατηγορίες περιλαμβάνουν λιγότερες λέξεις. Η κατηγορία «Υπεραγορά» 6 λέξεις, η κατηγορία «Φαγητά» 8 λέξεις, η κατηγορία «Ρούχα» 9 και η κατηγορία «Λαχανικά» 8 λέξεις υψηλής συχνότητας. Λέξεις αυτής της συχνότητας όπως «σκύλος», «τραπέζι», «πατάτες» φαίνεται πως χρησιμοποιούνται συχνότερα από τον Κυπριακό πληθυσμό. Οι λέξεις υψηλής συχνότητας ως σύνολο είχαν τις λιγότερες λέξεις από τις 3 κατηγορίες.

Στις λέξεις μέτριας συχνότητας, όπως καθορίστηκαν αυτές που καταγράφηκαν 4 έως 6 φορές, εντοπίζουμε λέξεις που χρησιμοποιούνται λίγο περισσότερο απ' ό τι οι λέξεις χαμηλής συχνότητας. Βλέπουμε πως οι κατηγορίες που εμφανίζουν τις περισσότερες λέξεις μεσαίας συχνότητας είναι η κατηγορία «Ζώα» η οποία περιλαμβάνει 20 λέξεις και η κατηγορία «Φαγητά» με 19 λέξεις. Παρουσιάζονται αρκετές λέξεις μέτριας συχνότητας και στην κατηγορία «Υπεραγορά» που περιλαμβάνει 17 λέξεις και στην κατηγορία «Αντικείμενα» 13 λέξεις, σε αντίθεση με τις κατηγορίες «Λαχανικά» και «Ρούχα» στις οποίες καταγράφηκαν 6 λέξεις. Λέξεις που παρουσιάζονται σε αυτή την κατηγορία συχνότητας φαίνεται πως είναι λέξεις που θα περιμέναμε να εντάσσονταν και στις λέξεις υψηλής αλλά και στις λέξεις χαμηλής συχνότητας (π.χ. μαξιλάρι, ρίγανη, καμηλοπάρδαλη).

Λέξεις χαμηλής συχνότητας ορίστηκαν όσες είχαν συχνότητα εμφάνισης από 1 έως 3. Παρατηρούμε λέξεις οι οποίες δεν χρησιμοποιούνται συχνά από τον Κυπριακό πληθυσμό. Η σημασιολογική κατηγορία στην οποία καταγράφηκαν οι περισσότερες λέξεις χαμηλής συχνότητας είναι «Αντικείμενα» με 97 λέξεις. Στη σημασιολογική

κατηγορία «Ζώα» παρατηρούνται 84 λέξεις, στην κατηγορία «Υπεραγορά» 83 λέξεις και στην κατηγορία «Λαχανικά» 73 λέξεις. Ωστόσο, η κατηγορία «Ρούχα» με 20 λέξεις και η κατηγορία «Λαχανικά» με 17 λέξεις παρουσιάζουν τις λιγότερες λέξεις χαμηλής συχνότητας. Λέξεις που εντάσσονται σε αυτή την κατηγορία όπως πεζούνη, καρότο, καπέλο θα αναμέναμε πως θα καταγράφονταν στις λέξεις μέτριας ή υψηλής συχνότητας. Οι λέξεις χαμηλής συχνότητας ως σύνολο παρατηρείται πως είναι περισσότερες από τις λέξεις μέτριας και υψηλής συχνότητας.

Βάσει των αποτελεσμάτων που προκύπτουν, διακρίνονται κάποιες διαφορές και ομοιότητες ανάμεσα στις σημασιολογικές κατηγορίες που παρουσιάστηκαν στους συμμετέχοντες. Στις σημασιολογικές κατηγορίες Ζώα, Υπεραγορά και Αντικείμενα, καταγράφηκαν περισσότερες λέξεις σε σύγκριση με τις υπόλοιπες κατηγορίες. Αυτό είναι πιθανόν να οφείλεται στο γεγονός ότι υπάρχουν πολύ περισσότερες λέξεις για τις σημασιολογικές κατηγορίες Ζώα, Υπεραγορά και Αντικείμενα από ότι στις κατηγορίες Ρούχα ή Λαχανικά. Ακόμη, παρατηρείται πως σε όλες τις σημασιολογικές κατηγορίες, οι λέξεις χαμηλής συχνότητας είναι περισσότερες. Επιπρόσθετα, βλέπουμε πως λέξεις που θα περιμέναμε να ανήκουν στις υψηλής συχνότητας λέξεις (π.χ.: αγρινό, πατάτες, παστοურμάς, μανιτάρια, αρνί, κολοκάσι), εν τέλει καταγράφονται στις χαμηλής συχνότητας λέξεις.

5 Συμπεράσματα

Όπως φαίνεται, η δημιουργία μιας λίστας συχνότητας για την Κυπριακή διάλεκτο μπορεί να αποτελέσει ένα σημαντικό εργαλείο για τον Κυπριακό πληθυσμό. Αρχικά, αυτό προκύπτει από το γεγονός πως μέσα από τα αποτελέσματα φαίνεται πως υπάρχουν αρκετές διαφορές συγκριτικά με τις λέξεις που συμπεριλαμβάνονται στις βάσεις δεδομένων που υπάρχουν ήδη για την Ελληνική γλώσσα. Στις ήδη υπάρχουσες βάσεις δεδομένων της Ελληνικής γλώσσας, οι λέξεις που τις αποτελούν έχουν παρθεί από δεδομένα που συλλέχθηκαν βάσει της γλώσσας που χρησιμοποιείται στην καθημερινότητα των Ελλήνων (Γούτσος, 2014). Όμως, στην Κύπρο δεν χρησιμοποιούνται οι ίδιες λέξεις, όροι και γλώσσα στην καθημερινότητα. Ακόμη, στις βάσεις δεδομένων που υπάρχουν για την Ελληνική γλώσσα, παρουσιάζονται οι λέξεις εντός προτάσεων (Γούτσος, 2014). Στην Κυπριακή διάλεκτο, όμως, υπάρχει διαφορετική σύνταξη, μορφολογία και φωνολογία. Εξαιτίας αυτών, λοιπόν, χρειάζεται να δημιουργηθεί μια ξεχωριστή λίστα συχνότητας και μια βάση δεδομένων με ειδικές κατηγορίες για την Κυπριακή διάλεκτο.

Η δημιουργία μιας λίστας συχνότητας και μετέπειτα μιας βάσης δεδομένων η οποία σχετίζεται με τη συχνότητα, μπορεί να οδηγήσει σε πολλά πλεονεκτήματα. Τα πλεονεκτήματα αυτά μπορούν να είναι μη κλινικά όπως να βοηθήσει ανθρώπους να μάθουν και να μελετήσουν την Κυπριακή διάλεκτο ή να αποτελέσει μια βάση ώστε να ενταχθεί η Κυπριακή διάλεκτος στην εκπαίδευση. Πιο σημαντικά πλεονεκτήματα, θα υπάρξουν στον κλινικό τομέα, αφού με τη δημιουργία μιας τέτοιας βάσης δεδομένων μπορούν να επωφεληθούν κλινικοί πληθυσμοί (π.χ.: άτομα με νόσο Αλτσχάιμερ). Αυτό μπορεί να συμβεί αφού βάσει του εντοπισμού της συχνότητας των λέξεων στον Κυπριακό πληθυσμό, μπορούν να δημιουργηθούν γνωστικά εργαλεία που θα στηρίζονται στη συχνότητα και θα βοηθούν στην αποκατάσταση ατόμων με γνωστικές ή άλλου είδους διαταραχές. Ένα παράδειγμα, είναι το γνωστικό εργαλείο ADAS – Cog (Alzheimer's Disease Assessment Scale-Cognitive Subscale), το οποίο βασίζεται στη συχνότητα των λέξεων και χρησιμοποιείται στην αξιολόγηση και αποκατάσταση ατόμων με νόσο Αλτσχάιμερ (Kueper, Speechley, & Montero-Odasso, 2018).

Η έρευνα αυτή προσφέρει σημαντικά αποτελέσματα αφού αποτελεί την πρώτη έρευνα αναφορικά με τη συχνότητα εμφάνισης των λέξεων στον Κυπριακό πληθυσμό.

Ωστόσο, υπάρχουν ορισμένοι περιορισμοί στην έρευνας μας. Αρχικά, το 90% των συμμετεχόντων είχαν ως τόπο διαμονής τη Λεμεσό και ως εκ τούτου το δείγμα δεν είναι αντιπροσωπευτικό του Κυπριακού πληθυσμού. Επίσης, οι συμμετέχοντες έπρεπε να είναι εξήντα ετών και άνω. Επομένως, δεν έλαβαν μέρος άτομα μικρότερης ηλικίας. Τέλος, το δείγμα της έρευνας ήταν μικρό. Λαμβάνοντας υπόψιν αυτούς τους περιορισμούς τα αποτελέσματα της μελέτης δεν είναι ασφαλή. Όμως, παρά τους περιορισμούς, η συγκεκριμένη λίστα συχνότητας των Κυπριακών λέξεων αποτελεί μια βάση για επόμενες έρευνες με παρόμοιο θέμα και σκοπό. Θα ήταν καλό να γίνουν μετέπειτα προσπάθειες δημιουργίας μιας λίστας συχνότητας λέξεων. Αυτό θα μπορούσε να επιτευχθεί μέσω ενός μεγαλύτερου δείγματος το οποίο θα αντιπροσωπεύει ηλικιακά όλες τις ομάδες του πληθυσμού αλλά και τις ευρύτερες περιοχές της Κύπρου. Κατ' αυτό τον τρόπο επόμενες μελέτες θα έχουν τη δυνατότητα να εξάγουν ασφαλέστερα συμπεράσματα για την πραγματική συχνότητα των λέξεων στον Κυπριακό πληθυσμό.

ΕΠΙΛΟΓΟΣ

Μέσα από την διαδικασία συλλογής δεδομένων, και την καταγραφή των λέξεων σύμφωνα με τη συχνότητα εμφάνισης φαίνεται πως υπάρχουν διαφορές ανάμεσα στις σημασιολογικές κατηγορίες σύμφωνα με τις λέξεις που εξέφρασαν οι συμμετέχοντες. Επίσης, εξάγεται το συμπέρασμα πως μια βάση δεδομένων η οποία θα έχει ως επίκεντρο τη συχνότητα εμφάνισης κάθε λέξης, μπορεί να αποτελέσει ένα σημαντικό εργαλείο για τον Κυπριακό πληθυσμό. Μέσω μιας βάσης δεδομένων η οποία θα βασίζεται στην συχνότητα, μπορούν να επωφεληθούν πολλά άτομα και ειδικοί στην Κύπρο.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Avdyli, S. R., & Cuetos, S. F. (2013). SUBTLEX-AL: Albanian word frequencies based on film subtitles. *ILIRIA International Review*, 3(1), 285-292.
- Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, 55(2), 290-313.
- Badham, S. P., Whitney, C., Sanghera, S., & Maylor, E. A. (2017). Word frequency influences on the list length effect and associative memory in young and older adults. *Memory*, 25(6), 816-830
- Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human perception and performance*, 10(3), 340.
- Balota, D. A., Burgess, G. C., Cortese, M. J., & Adams, D. R. (2002). The word-frequency mirror effect in young, old, and early-stage Alzheimer's disease: Evidence for two processes in episodic recognition performance. *Journal of Memory and Language*, 46(1), 199-226.
- Brébion, G., David, A. S., Bressan, R. A., & Pilowsky, L. S. (2005). Word frequency effects on free recall and recognition in patients with schizophrenia. *Journal of Psychiatric Research*, 39(2), 215-222.
- Brybaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect. *Experimental psychology*.
- Brybaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, 27(1), 45-50
- Brybaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods*, 41(4), 977-990.

- Brysbaert, M., Keuleers, E., & New, B. (2011). Assessing the usefulness of Google Books' word frequencies for psycholinguistic research on word processing. *Frontiers in Psychology, 2*, 27.
- Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior research methods, 44*(4), 991-997.
- Brysbaert, M., Keuleers, E., Stevens, M., Van der Haegen, L., Verma, A., Callens, M., ... & Princen, M. Subtitle word frequencies for Spanish: SUBTLEX-ESP.
- Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior research methods, 44*(4), 991-997.
- Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language variation and change, 14*(3), 261-290.
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS one, 5*(6).
- Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2012). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica, 33*(2), 133-143.
- Cockrell, J. R., & Folstein, M. F. (2002). Mini-mental state examination. *Principles and practice of geriatric psychiatry, 140-141*.
- Dick, J. P., Guiloff, R. J., Stewart, A., Blackstock, J., Bielawska, C., Paul, E. A., & Marsden, C. D. (1984). Mini-mental state examination in neurological patients. *Journal of Neurology, Neurosurgery & Psychiatry, 47*(5), 496-499.
- Dimitropoulou, M., Duñabeitia, J. A., Avilés, A., Corral, J., & Carreiras, M. (2010). Subtitle-based word frequencies as the best estimate of reading behavior: The case of Greek. *Frontiers in psychology, 1*, 218.
- de Groot, A. M. (2006). Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Language Learning, 56*(3), 463-506.

- Ellis, A. W., Miller, D., & Sin, G. (1983). Wernicke's aphasia and normal language processing: A case study in cognitive neuropsychology. *Cognition*, 15(1-3), 111-144.
- Galea, M., & Woodward, M. (2005). Mini-mental state examination (MMSE). *Australian Journal of Physiotherapy*, 51(3), 198.
- Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, 2(1), 21.
- Hatzigeorgiu, N., Gavrilidou, M., Piperidis, S., Carayannis, G., Papakostopoulou, A., Spiliotopoulou, A., ... & Demiros, I. (2000, May). Design and Implementation of the Online ILSP Greek Corpus. In *LREC*.
- Hayakawa, S., Ning, S., & Marian, V. (2020). From Klingon to Colbertian: Using artificial languages to study word learning. *Bilingualism: Language and Cognition*, 23(1), 74-80.
- Hill, K. J., & Romich, B. A. (2000). AAC core vocabulary analysis: Tools for clinical use. In *Proceedings of the RESNA 2000 Annual Conference* (Vol. 3, pp. 67-69).
- Hoffman, P., Rogers, T. T., & Lambon Ralph, M. A. (2011)
- Howes, D. (1964). Application of the word-frequency concept to aphasia. In *Disorders of language* (pp. 47-75). Churchill London.
- Huron, C., Danion, J.-M., Giacomoni, F., Grangé, D., Robert, P., & Rizzo, L. (1995). Impairment of recognition memory with, but not without, conscious recollection in schizophrenia. *American Journal of Psychiatry*, 152, 1737-1742.
- Jescheniak, J. D., & Levelt, W. J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 824.

- Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual cognition*, 13(7-8), 789-845.
- Jiménez González, J. E., & Hernández Valle, I. (2000). Word identification and reading disorders in the Spanish language. *Journal of Learning disabilities*, 33(1), 44-60.
- Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior research methods*, 42(3), 643-650.
- Ktori, M., van Heuven, W. J., & Pitchford, N. J. (2008). GreekLex: A lexical database of Modern Greek. *Behavior Research Methods*, 40(3), 773-783.
- Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Dartmouth Publishing Group.
- Kueper, J. K., Speechley, M., & Montero-Odasso, M. (2018). The Alzheimer's disease assessment scale–cognitive subscale (ADAS-Cog): modifications and responsiveness in pre-dementia populations. a narrative review. *Journal of Alzheimer's Disease*, 63(2), 423-444.
- Kyparissiadis, A., van Heuven, W. J., Pitchford, N. J., & Ledgeway, T. (2017). GreekLex 2: A comprehensive lexical database with part-of-speech, syllabic, phonological, and stress information. *PloS one*, 12(2).
- Lancu, I., & Olmer, A. (2006). The minimal state examination--an up-to-date review. *Harefuah*, 145(9), 687-90.
- LIU, C., & SLOANE, Z. (2006). Developing a Core Vocabulary for a Mandarin Chinese AAC System Using Word Frequency Data. *International Journal of Computer Processing of Languages*, 19(04), 285–300.

- Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior research methods, instruments, & computers*, 28(2), 203-208.
- M. A. (2007). *Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. Quarterly Journal of Experimental Psychology*, 60(1), 32–42
- Marian V (2017) Orthographic and phonological neighborhood databases across multiple languages. *Written Language & Literacy* 20, 6–26
- Mandera, P., Keuleers, E., Wodniecka, Z., & Brysbaert, M. (2015). Subtlex-pl: subtitle-based word frequency estimates for Polish. *Behavior research methods*, 47(2), 471-483.
- Mikros, G., Hatzigeorgiu, N., & Carayannis, G. (2005). Basic quantitative characteristics of the Modern Greek language using the Hellenic National Corpus. *Journal of Quantitative Linguistics*, 12(2-3), 167-184.
- Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they?. *Journal of Experimental Psychology: General*, 118(1), 43.
- Nation, K., Marshall, C. M., & Snowling, M. J. (2001). Phonological and semantic contributions to children's picture naming skill: Evidence from children with developmental reading disorders. *Language and Cognitive Processes*, 16(2-3), 241-259.
- Newman German, D. J. (1979). Word-finding skills in children with learning disabilities. *Journal of Learning Disabilities*, 12(3), 176-181.
- New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied psycholinguistics*, 28(4), 661-677.

- Nikolaev, A., Ashaie, S., Hallikainen, M., Hänninen, T., Higby, E., Hyun, J., ...
Soininen, H. (2018). Effects of morphological family on word recognition in normal aging, mild cognitive impairment, and Alzheimer's disease. *Cortex*. doi:10.1016/j.cortex.2018.10.028
- Nozari, N., Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2010). Naming and repetition in aphasia: Steps, routes, and frequency effects. *Journal of memory and language*, 63(4), 541-559.
- Pavlou, P., & Papapavlou, A. (2004). Issues of dialect use in education from the Greek Cypriot perspective. *International Journal of Applied Linguistics*, 14(2), 243-258.
- Semantic diversity accounts for the "missing" word frequency effect in stroke aphasia: Insights using a novel method to quantify contextual variability in meaning. *Journal of Cognitive Neuroscience*, 23(9), 2432-2446.
- Soares, A. P., Machado, J., Costa, A., Iriarte, Á., Simões, A., de Almeida, J. J., ... & Perea, M. (2015). On the advantages of word frequency and contextual diversity measures extracted from subtitles: The case of Portuguese. *The Quarterly Journal of Experimental Psychology*, 68(4), 680-696.
- Sophocleous, A., & Wilks, C. (2010). Standard Modern Greek and Greek-Cypriot dialect in kindergarten classroom interaction: teachers' and learners' language attitudes and language use. *Language, culture and curriculum*, 23(1), 51-69.
- Terzopoulos, A. R., Duncan, L. G., Wilson, M. A., Niolaki, G. Z., & Masterson, J. (2017). HelexKids: A word frequency database for Greek and Cypriot primary school children. *Behavior research methods*, 49(1), 83-96.
- Terzopoulos, A. R. The importance of psycholinguistic databases for research with children: The case of HelexKids, a Greek word database for children. Using crowdfunding to support your PhD research The application of a qualitative approach to adult attachment research: A short report Pregnant women, mothers, mother and, 37

- Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, 67(6), 1176-1190.
- Word frequency influences on the list length effect and associative memory in young and older adults. *Memory*, 25(6), 816-830.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of memory and language*, 46(3), 441-517.
- Γούτσος, Δ. (2014). Για μια γραμματική του ελληνικού προφορικού λόγου. *Μελέτες αφιερωμένες στην Ομότιμη Καθηγήτρια ΑΠΘ Άννα Αναστασιάδη-Συμεωνίδη*, 211.
- Μπουρονίκου, Φ., Παπαντωνίου, Γ., Μωραΐτου, Δ., & Σαρρής, Δ. (2014). Η χρήση της δοκιμασίας Mini-Mental State Examination (MMSE) σε άτομα παιδικής και εφηβικής ηλικίας: Βιβλιογραφική ανασκόπηση ερευνητικών δεδομένων. *Επιστημονική Επετηρίδα Παιδαγωγικού Τμήματος Νηπιαγωγών Πανεπιστημίου Ιωαννίνων*, 7, 122-180.

ΠΑΡΑΡΤΗΜΑ Ι

Λίστα συχνότητας λέξεων σημασιολογικού περιεχομένου



Semantic Word Frequency List Project

«Θα σας αναφέρω μια σημασιολογική κατηγορία και θέλω να προσπαθήσετε να μου πείτε όσες περισσότερες λέξεις πιστεύετε ότι ταιριάζουν στη κατηγορία αυτή». Αφού ο συμμετέχοντας ανακαλέσει ελεύθερα όσες περισσότερες λέξεις μπορεί, τότε του δίνονται οι υπόλοιπες κατηγορίες. «Τώρα, θα ήθελα να μου πείτε όσα περισσότερα κατοικίδια ΖΩΑ μπορείτε να σκεφτείτε».

Ζώα							
	Ελεύθερη ανάκληση	Κατοικίδια	Σαρκοφάγα	Ψάρια	Πουλιά	Έντομα	Φάρμας
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							

Πράγματα που αγοράζεις στην Υπεραγορά				
	Ελεύθερη ανάκληση			
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				

Φαγητά/Πράγματα που τρώμε				
	Ελεύθερη ανάκληση	Γλυκά	Μπαχαρικά	Τυριά
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				

Λαχανικά				
	Ελεύθερη ανάπτυξη	Σαλάτας	Πράσινα	Με φύλλα
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				



Ρούχα					
	Ελεύθερη ανάκλιση	Γυναικεία	Αντρικά	Καλοκαιρινά	Χειμερινά
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					

Αντικείμενα/Πράγματα					
	Ελεύθερη ανάκληση	Που βρίσκεις στη κουζίνα	Που βρίσκεις σε ένα γραφείο	Έπιπλα	Εργαλεία
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					

Ποτά/Ροφήματα				
	Ελεύθερη ανάσπληση	Οινοπνευματώδη	Ροφήματα/ποτά που πίνουμε το χειμώνα	Ροφήματα/ποτά που πίνουμε το καλοκαίρι
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				



	Μουσικά Όργανα	Οχήματα	Μέρη/Κτήρια	Λαυλούδια	Χώροι σπιτιού
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					

ΠΑΡΑΡΤΗΜΑ ΙΙ

Πίνακες συχνότητας συλλεγόμενων λέξεων ανά σημασιολογική κατηγορία:

Πίνακας 1: Συχνότητα λέξεων σημασιολογικής κατηγορίας Ζώα

Πίνακας Συχνότητας Λέξεων: ΖΩΑ		
Χαμηλή	Μέτρια	Υψηλή
αγρινό (1)	λεοπάρδαλη (4)	κατσίκα (7)
πεταλούδα (1)	φαγκρί (4)	κουνέλι (7)
ταύρος (1)	κροκόδειλος (4)	τσιπούρα (7)
πάπια (1)	πετεινός (4)	πέρδικα (7)
χήνα (1)	περιστέρι (4)	χταπόδι (8)
καρακάξα (1)	κουκουβάγια (4)	μούγια (8)
γαρίδα (1)	αμπελοπούλι (4)	άλογο (8)
σπουργίτι (1)	σκαθάρι (4)	σκνίπα (8)
μουλάρι (1)	καναρίνι (4)	παπαγάλος (8)
ρινόκερος (1)	τζίκλα (5)	αγελάδα (8)
πάνθηρας (1)	πίθηκος (5)	κότα (8)
αρνί (1)	σουπιά (5)	ελέφαντας (9)
χοίρος (1)	μέλισσα (5)	γάιδαρος (9)
αλιγάτορας (1)	πέστροφα (5)	φίδι (9)
αίγα (1)	αλεπού (5)	κατσαρίδα (9)
κριάρι (1)	γάλος (5)	γάτα (10)
πούμα (1)	τίγρης (6)	πρόβατο (10)
πύθωνας (1)	καμηλοπάρδαλη (6)	σκύλος (10)
αγριογούρουνο (1)	λυθρίνι (6)	λιοντάρι (10)
χιμπατζής (1)	γουρούνι (6)	
αγριόχηνα (1)		
ουρακοτάγκος (1)		
κοριδαλός (1)		
τρασχύλα (1)		
καρδερίνα (1)		
ματσιόνα (1)		
φαραόνα (1)		
τσούρα (1)		
τσακάλι (1)		
κόμπρα (1)		
ύαινα (1)		

λυκόσκυλο (1)		
βόρτος (1)		
ακρίδα (1)		
δελφίνι (1)		
χαμολιός (1)		
σουρικάτα (1)		
σολωμός (1)		
χέλι (1)		
τσαλαπετεινός (1)		
κουρκουτάς (1)		
πόνι (1)		
ιππόκαμπος (1)		
λαμούριος (1)		
ζέβρα (1)		
τσούχτρα (1)		
μαρτάκι (1)		
λαγός (1)		
ψάρι (1)		
τάρανδος (1)		
αράχνη (2)		
χελώνα (2)		
ποντικός (2)		
τίγρης (2)		
τσίτα (2)		
χελιδόνι (2)		
λύκος (2)		
μαϊμού (2)		
φάλαινα (2)		
αράχνη (2)		
καρχαρίας (2)		
φιλικουτούνι (2)		
στρούθος (2)		
νυχτερίδα (2)		
γύπας (2)		
σαρανταποδαρούσα (2)		
χρυσόψαρο (2)		
βόας (2)		
σκαντζόχοιρος (2)		

καμήλα (2)		
κούλα (2)		
χάμστερ (2)		
ιπποπόταμος (2)		
πολική αρκούδα (2)		
ελάφι (3)		
γορίλλας (3)		
γόπα (3)		
μπαρμπούνι (3)		
κοράζινος (3)		
σφήκας (3)		
λίμπουρος (3)		
πεζούνι (3)		
λαβράκι (3)		
αετός (3)		

Πίνακας 2: Συχνότητα λέξεων σημασιολογικής κατηγορίας Υπεραγορά

Πίνακας Συχνότητας Λέξεων: ΥΠΕΡΑΓΟΡΑ		
Χαμηλή	Μέτρια	Υψηλή
αναψυκτικά (1)	λάδι (4)	πατάτες (7)
απολυμαντικά (1)	τσάι (4)	ζάχαρη (7)
ρολά (1)	ρεβίθια (4)	γάλα (7)
καθαριστικά τουαλέτας (1)	μέλι (4)	καφές (8)
κοτόσουπες (1)	μαρμελάδα (4)	κρεμμύδια (8)
ξύδι (1)	χυμός (4)	βούτυρος (8)
μπύρες (1)	νερό (4)	
ούισκι (1)	ρύζι (5)	
κοκα – κόλα (1)	λούντζα (4)	
φυστίκελαιο (1)	μαρούλι (5)	
χαρουπόμελο (1)	κρέας (5)	
κολοκάσι (1)	σοκολάτες (5)	
σέλινο (1)	μακαρόνια (5)	
μαλακτικό ρούχων (1)	ντομάτα (6)	
μανιτάρια (1)	αγγούρι (6)	
σπίρτα (1)	φασόλια (6)	
πίττες (1)	ψωμί (6)	

σεφταλιά (1)		
καρβουνάκια (1)		
χλωρίνη (1)		
ζαρζαβατικά (1)		
κατεψυγμένα (1)		
ρόκα (1)		
φράουλες (1)		
μήλα (1)		
τριαντάφυλλο (1)		
πορτοκαλάδα (1)		
λεμονάδα (1)		
κακάο (1)		
λεμόνια (1)		
μπριζόλες (1)		
κρουασάν (1)		
παγωτό (1)		
μπανάνες (1)		
πεπόνι (1)		
μανταρίνια (1)		
καρπούζι (1)		
τυρόπιτα (1)		
αλάτι (1)		
ραβιόλες (1)		
τσίπς (1)		
γατοτροφή (1)		
παιχνίδια (1)		
φέτα (1)		
αγκινάρες (1)		
νέσκουικ (1)		
κρέμα γάλακτος (1)		
μπρόκολο (1)		
καθαριστικά (1)		
μαλακτικό ρούχων (2)		
χαμ (2)		
σαπούνι (2)		
φακές (2)		
δυόσμος (2)		
λάχανα (2)		

κειμάς (2)		
μαγιά (2)		
αλεύρι (2)		
πιπεριές (2)		
κραμπί (2)		
πορτοκάλια (2)		
πουρκούρι (2)		
χόρτα (2)		
πόλιπιφ (2)		
παξιμάδια (2)		
καρότα (2)		
μπέικον (2)		
τραχανάς (2)		
νέσκαφε (2)		
υγρό σαπούνι (2)		
μαιντανός (2)		
κουκιά (3)		
κριθαράκι (3)		
γαλοπούλα (3)		
κόφλεξ (3)		
μπισκότα (3)		
κοτόπουλο (3)		
κάρβουνα (3)		
γιαούρτι (3)		
σαλάμι (3)		
λουβί (3)		
σαμπουάν (3)		
σκόνη πλυσίματος (3)		

Πίνακας 3: Συχνότητα λέξεων σημασιολογικής κατηγορίας Φαγητά

Πίνακας Συχνότητας Λέξεων: ΦΑΓΗΤΑ		
Χαμηλή	Μέτρια	Υψηλή
ψτουρμάς (1)	φακή (4)	ψητό (7)
αναρή(1)	γεμιστά (4)	μακαρόνια (7)
σπαγγέτι (1)	αρτισιά (4)	χαλούμι (7)

αυγολέμονη (1)	μακαρόνια με κειμά (4)	κασκαβάλι (8)
χορτόσουπα (1)	πατάτες τηγανιτές (4)	πιπέρι (8)
τραχανάς (1)	λουκάνικα (4)	κουπέπια (8)
μανιτάρια (1)	γλυκό κιτρόμηλο (4)	φασόλες (9)
μουσακάς (1)	κανέλα (5)	σούβλα (9)
κουνέλι στιφάδο (1)	μακαρόνια του φούρνου (5)	
γαρίδες (1)	πακλαβάς (5)	
πουρέκκια (1)	ρίγανη (5)	
πόμπες (1)	τούρτα (5)	
μους σοκολάτα (1)	γλυκά κουταλιού (5)	
σιάμαλι (1)	φέτα (5)	
καττιμέρι (1)	γαλακτομπούρεκο (6)	
πουτίγκα (1)	λουβί (6)	
πίττες της σάτζις (1)	σουβλάκια (6)	
κουραμπιέδες (1)	δάφνη (6)	
μελομακάρονα (1)	αναρή (6)	
κρίστμας κέικ (1)		
φλογέρες (1)		
κρέμες (1)		
πιπέρι μαύρο (1)		
γλυκάνισος (1)		
γαρύφαλλο (1)		
μαστίχα (1)		
μαχλέπι (1)		
πάπρικα (1)		
τσίλι (1)		
λεμονάτο (1)		
ψητό (1)		
πατάτες με αυγά (1)		
γλυκό χρυσόμηλο (1)		
γλυκό μόσφυλο (1)		
γλυκό μήλο (1)		
γλυκό σκόρδος (1)		
κίμινο (1)		
σκόρδο (1)		
κρεμμύδι (1)		
κρέας (1)		
όσπρια (1)		

χόρτα βραστά (1)		
ψάρι (1)		
μηλόπιτα (1)		
χαλλουμοτή (1)		
φρούτα (1)		
κέικ (1)		
κεφαλάκια (1)		
πατάτες (1)		
ραβιόλες (1)		
κριθαράκι (1)		
μπριζόλες (1)		
τιραμισού (1)		
κουκιά με λάχανα (1)		
λουβί με λάχανα (1)		
λαζάνια (1)		
παγωτό (1)		
καραμελέ (1)		
ρολό (2)		
ρόστο (2)		
γιαούρτι (2)		
κούννες (2)		
σοκολάτες (2)		
κανέλα (2)		
τζιέλι (2)		
προφιτερόλ (2)		
δούκισσα (2)		
σπανακόρυζο (2)		
ρύζι (2)		
πουρκούρι (2)		
πατάτες του φούρνου (2)		
ελιοτή (2)		
στιφάδο (2)		
θυμάρι (2)		
μπιζέλια (2)		
τυρόπιτα (2)		
ταχινόπιτα (2)		
κατείφι (2)		
κοτολέτες (2)		

σούπα (2)		
κεφαλοτύρι (2)		
γλυκό κρεμμύδι (2)		
γλυκό πορτοκάλι (2)		
γλυκό πατίχα (2)		
μπιφτέκια (2)		
κεφτέδες (2)		
φασολάκι (2)		
αγρέλια με αυγά (2)		
γλυκό περγαμόντο (2)		
γλυκό αχλάδι (2)		
κουκιά (2)		
σιεφταλιές (3)		
κολοκάσι (3)		
ρεβύθια (3)		
γλυκό καρυδάκι (3)		
τυρί (3)		
κοτόπουλο (3)		
πατάτες (3)		
χαλουβάς (3)		
γλυκό κεράζι (3)		
δύοσμος (3)		
τσισκέικ (3)		
ψωμί (3)		
σοκολατίνα (3)		

Πίνακας 4: Συχνότητα λέξεων σημασιολογικής κατηγορίας Λαχανικά

Πίνακας Συχνότητας Λέξεων: ΛΑΧΑΝΙΚΑ		
Χαμηλή	Μέτρια	Υψηλή
κάρταμα (1)	δύοσμος (4)	λάχανα (7)
κουτρούβι (1)	πιπεριές (4)	σέλινο (7)
κρεμμύδι (1)	γλυστιρίδα (5)	σπανάχι (7)

λαχανάκια Βρυξελλών (1)	κόλιαντρος (5)	αγγούρι (8)
λαψάνες (1)	μαϊντανός (6)	ντομάτα (8)
στουφούθκια (1)	μπρόκολο (6)	μαρούλι (9)
χόρτα (1)		κραμπί (10)
ραδίκια (1)		ρόκα (10)
καπάρι (2)		
κουλούμπρα (2)		
λουβάνα (2)		
κρεμμύδι φρέσκο (3)		
ραπανάκι (3)		
καρότο (3)		
κουνουπίδι (3)		
κρεμμύδι (3)		
λάχανα (3)		

Πίνακας 5: Συχνότητα λέξεων σημασιολογικής κατηγορίας Ρούχα

Πίνακας Συχνότητας Λέξεων: ΡΟΥΧΑ		
Χαμηλή	Μέτρια	Υψηλή
βερμούδα (1)	μπλούζα (4)	ζακέτα (7)
γιλέκο (1)	παλτό (4)	πουκάμισο (7)
γραβάτα (1)	μαγιό (5)	τρικό (7)
κλατσό (1)	παπούτσια (5)	φούστα (8)
μπότες (1)	παντόφλες (6)	φανελά (8)
πιζιάμες (1)	εσώρουχα (6)	φουστάνι (9)
πόλο (1)		σακάκι (9)
σουτιεν (1)		παντελόνι (9)
ρόπα (1)		κλάτσες (10)
πουλόβερ (1)		
φόρμες (2)		
τελαντωτή (2)		
καπαρτίνα (2)		
κουστούμι (2)		
σκούφος (2)		
σώβρακο (3)		
ταγιέρ (3)		
σιάλπα (3)		

καπέλο (3)		
γάντια (3)		

Πίνακας 6: Συχνότητα λέξεων σημασιολογικής κατηγορίας Αντικείμενα

Πίνακας Συχνότητας Λέξεων: ANTIKEIMENA		
Χαμηλή	Μέτρια	Υψηλή
φωτιστικό (1)	τηλέφωνο (4)	πέννα (7)
φλόκκος (1)	κάβουρας (4)	κουτάλα (7)
σκούπα (1)	πένσα (4)	καναπές (7)
ρούχα (1)	τσάντα (4)	ποτήρι (8)
γυαλιά (1)	κομοδίνο (4)	κουτάλι (8)
τρυπάνι (1)	μαξιλάρι (4)	μαχαίρι (8)
φάκελος (1)	μέτρο (4)	πιρούνι (8)
ανεμιστήρας (1)	σφυρίλιος (5)	κομπιούτερ (8)
πορτοφόλι (1)	φάιλ (5)	τηγάνι (8)
κορνίζα (1)	τηλεόραση (6)	κρεβάτι (8)
παντόφλες (1)	βιβλίο (6)	μολύβι (8)
ζυγαριά (1)	ψυγείο (6)	κατσαρόλα (9)
κάλαθος (1)	γραφείο (6)	καρέκλες (9)
πριόνι (1)		κατσαβίδι (9)
σιφονιέρα (1)		σφυρί (9)
ξύστρα (1)		τραπέζι (10)
βραστήρας (1)		
σεντόνια (1)		
πάπλωμα (1)		
αμπαζούρ (1)		
σημειωματάριο (1)		
μολυβοθήκες (1)		
σπόντες (1)		
κουτάλα (1)		
τραπεζομάντηλο (1)		
μαδράτζι (1)		
χαρτί κουζίνας (1)		
γυαλιά ηλίου (1)		
καπέλο (1)		
λάμπα (1)		
φωτογραφοθήκη (1)		

αλατιέρα (1)		
εργαλειοθήκη (1)		
παράθυρο (1)		
κάγκελο (1)		
ραδιόφωνο (1)		
κιβώτια (1)		
παιχνίδια (1)		
σκάλα (1)		
μάτσα (1)		
ηλεκτροκόλληση (1)		
κοντρόλ (1)		
ξύλο (1)		
πλάστης (1)		
κουταλάκι (1)		
σουηδικό κλειδί (1)		
σβανάς (1)		
σκεπάρνι (1)		
μιστρί (1)		
κλαδευτήρι (1)		
χαλί (1)		
χαρτομάντιλα (1)		
ψωμοθήκη (1)		
μπρίκι (1)		
λιβέρι (1)		
χαρτόνι (1)		
τρύφτης (1)		
κλωστές (1)		
κάδρο (1)		
πιατέλες (2)		
τραπανάκι (2)		
σόμπα (2)		
κουρτίνες (2)		
μπουκάλα (2)		
ραπτομηχανή (2)		
βάζο (2)		
κασετίνα (2)		
φλιτζάνι (2)		
τοστιέρα (2)		

μίξερ (2)		
πέννα (2)		
πίνακας (2)		
τσεκκούρι (2)		
μαντιλιά (2)		
φτυάρι (2)		
χτένα (2)		
βιβλιοθήκη (2)		
βούρνα (2)		
καθρέφτης (3)		
συρτάρια (3)		
φούρνος (3)		
κούππα (3)		
πάγκος (3)		
γκάζι (3)		
σιεπάρνι (3)		
συρταριέρα (3)		
αρμάρι (3)		
σίδερο (3)		
πόρτα (3)		
σιγάτσα (3)		
τσάππα (3)		
κούσπος (3)		
τετράδιο (3)		
σινί (3)		
κουβέρτα (3)		
ψαλίδι (3)		
καρφίτσες (3)		