

Ένα πολυτροπικό σύστημα αρχειοθέτησης και διαχείρισης τεκμηρίων για τη γλώσσα των Ελλήνων μεταναστών του Καναδά

Χαράλαμπος Τσιμπούρης, Σταύρος Μπόμπολας, Χρήστος Παπαναγιώτου,
Γιώργος Χαιρετάκης, Βάσω Αλεξέλλη & Αγγελική Ράλλη
Πανεπιστήμιο Πατρών

Περίληψη

Στην παρούσα εργασία παρουσιάζονται το ψηφιακό αποθετήριο (στο εξής ΨΑ) και η βάση δεδομένων (στο εξής ΒΔ) που αναπτύχθηκαν στο πλαίσιο του ερευνητικού προγράμματος “Immigration and Language in Canada. Greeks and Greek-Canadians” (στο εξής ImmiGrec), με σκοπό τη δημιουργία και διαχείριση ενός σώματος τεκμηρίων για τη γλώσσα της πρώτης γενιάς Ελλήνων μεταναστών στον Καναδά (μέσα '40 - τέλος '70). Πιο συγκεκριμένα, το σώμα τεκμηρίων συνίσταται από πρωτογενή και επεξεργασμένα δεδομένα που παρουσιάζουν διάφορες πτυχές της ζωής των Ελλήνων του Καναδά, δίνοντας έμφαση σε πληροφορίες αναφορικά με τον τρόπο με τον οποίο επηρεάστηκε η μητρική γλώσσα, τα σημεία αναχώρησης και άφιξης, τη διαδικασία και τις δυσκολίες ενσωμάτωσης στο καναδικό περιβάλλον, την οργάνωση των ελληνικών κοινοτήτων σε καναδικές πόλεις και, τέλος, την πολιτιστική τους ταυτότητα και τη συλλογική μνήμη. Το υλικό καταχωρείται σε ένα ΨΑ (της πλατφόρμας ownCloud) και είναι οργανωμένο σε μια ηλεκτρονική ΒΔ κατασκευασμένη με το Drupal 7 CMS (v.7). Και οι δύο υπηρεσίες παρέχονται μέσω της ίδιας τεχνολογικής υποδομής και γίνεται χρήση της στοίβας εφαρμογών *Linux/Apache/MySQL/PHP* (γνωστή ως LAMP).

Λέξεις-κλειδιά: σώμα τεκμηρίων, βάση δεδομένων, ψηφιακό αποθετήριο, ελληνική μετανάστευση, Καναδάς.

1. Εισαγωγή

Η δημιουργία και διαχείριση ενός αντιπροσωπευτικού σώματος γλωσσικών δεδομένων αποτέλεσε πρόκληση ήδη από την αρχή της μελέτης των γλωσσικών ποικιλιών με τον Georg Wenker στα τέλη του 19^{ου} αιώνα. Είναι χαρακτηριστικό ότι ο Wenker, για να μπορέσει να διαχειριστεί τον τεράστιο όγκο δεδομένων που συνέλεξε από 45.000 ερωτηματολόγια για ποικιλίες της Γερμανικής, περιορίστηκε μόνο σε αυτά που προήλθαν από την Κεντρική και Βόρεια Γερμανία, ενώ χρειάστηκαν δέκα χρόνια για την παρουσίαση των πρώτων αποτελεσμάτων και πάνω από σαράντα για την ολοκλήρωση της μελέτης. Από τότε, για κάθε νέα γλωσσική μελέτη μεγάλης κλίμακας, οι γλωσσολόγοι χρησιμοποιούσαν όλο και περισσότερες πληροφορίες και δεδομένα, ο όγκος των οποίων αυξήθηκε ραγδαία την εποχή που χρησιμοποιήθηκαν τα φορητά κασετόφωνα για την ηχογράφηση των διαλεκτόφωνων (Chambers & Trudgill 1980).

Σήμερα, οι νέες τεχνολογίες που προέκυψαν με τη χρήση των ηλεκτρονικών υπολογιστών και τη ραγδαία ανάπτυξη των Ψηφιακών Ανθρωπιστικών Επιστημών (Digital Humanities) έχουν διευκολύνει τη διαχείριση του τεράστιου όγκου δεδομένων που μπορεί να συλλεγεί από μία γλωσσική μελέτη μεγάλης κλίμακας. Στη γραμμή αυτή, η δημιουργία ψηφιακών βάσεων γλωσσικών δεδομένων αποτελεί καταλυτικό εργαλείο, γεγονός που αποδεικνύεται από το πλήθος των ψηφιακών βάσεων για πολλές γλώσσες (και γλωσσικές ποικιλίες) του κόσμου. Πολλές από αυτές συνιστούν σημεία αναφοράς

λόγω του γλωσσικού περιεχομένου τους, αλλά και της χρηστικότητας τους, καθώς επιτρέπουν τη γρήγορη και ακριβή πλοήγηση μέσα σε έναν τεράστιο όγκο δεδομένων.

Ενδεικτικά, αξίζει να σημειωθούν οι ακόλουθες βάσεις γλωσσικών δεδομένων. *DynaSAND corpus*¹ (Dynamic Syntactic Atlas of Dutch Dialects): αποτελεί ένα διαδικτυακό εργαλείο για την επεξεργασία των συντακτικών ποικιλιών της Ολλανδικής (βλ. Barbiers *et al.* 2006· Kunst & Wesseling 2010)· *IDEA*² (International Dialects of English Archive, βλ. Meier *et al.* 1998), για τις ποικιλίες της Αγγλικής· *ONZE corpus*³ (Origins of New Zealand English, βλ. Bayard 2000)· *FRED corpus*⁴ (Freiburg English Dialect Corpus, βλ. Anderwald & Wagner 2007)· *NECTE corpus*⁵ (Newcastle Electronic Corpus of Tyneside English, βλ. Beal *et al.* 2007)· *SCOTS corpus*⁶ (Scottish Corpus Of Text & Speech, βλ. Anderson & Corbett 2009)· *ScaDiaSyn*⁷ (Scandinavian Dialect Syntax-Nordic Dialect Corpus and Syntax Database, βλ. Bondi *et al.* 2014)· *TGDP*⁸ (Texas German Dialect Project, βλ. Boas 2002)· *SADS*⁹ (Swiss German Dialects, βλ. Glaser 2013).

Μόλις τα τελευταία χρόνια έχει υπάρξει μέριμνα για τη δημιουργία ψηφιακών βάσεων με δεδομένα από γλωσσικές ποικιλίες της Ελληνικής. Μία γενική ΒΔ είναι η *Gree.D.* (Greek Dialects), η οποία αποτελεί την πρώτη ελληνική πολυτροπική διαλεκτική ΒΔ (Ράλλη *et al.* 2010) με 500 περίπου ώρες υλικό διαλεκτικού προφορικού λόγου. Απαρτίζεται από 15 μικρότερες βάσεις και εμπλουτίζεται συνεχώς με νέα δεδομένα, τα οποία συγκεντρώνονται στο πλαίσιο των ερευνητικών προγραμμάτων που εκπονούνται στο *Εργαστήριο Νεοελληνικών Διαλέκτων*¹⁰ του Πανεπιστημίου Πατρών. Μία ακόμα στοχευμένη διαλεκτική πολυτροπική ΒΔ είναι η *AMiGreDB*¹¹ που δημιουργήθηκε στο πλαίσιο του ερευνητικού προγράμματος «ΘΑΛΗΣ: Πόντος, Καππαδοκία, Αίβαλι: στα χνάρια της Μικρασιατικής Ελληνικής» (Galiotou *et al.* 2014· Ράλλη & Μπόμπολας 2015). Η βάση αυτή περιλαμβάνει ένα ψηφιακό σώμα γραπτών πηγών και από τις τρεις γλωσσικές ποικιλίες έκτασης 2.000.000 λέξεων, καθώς και προφορικά δεδομένα 180 ωρών περίπου (60 ώρες ανά διάλεκτο). Επιπλέον, ενσωματώνει ένα εύρος εργαλείων πλοήγησης, διαχείρισης, επεξεργασίας και αναζήτησης για τα γλωσσικά δεδομένα.

Το ΨΑ και η ΒΔ που δημιουργήθηκαν στο πλαίσιο του προγράμματος *ImmiGrec* αποτελούν την πρώτη ολοκληρωμένη και συστηματική προσπάθεια να καταγραφεί η γλώσσα των Ελλήνων του Καναδά και περιέχουν πλήθος πρωτογενών τεκμηρίων –συνεντεύξεις, φωτογραφίες και διάφορα άλλα έγγραφα αρχαιικού τύπου– με τη μορφή αρχείων ήχου, εικόνας και βίντεο. Πρόκειται για μια συνεχώς διευρυνόμενη ΒΔ, που εξασφαλίζει ότι οι αφηγήσεις των Ελλήνων μεταναστών είναι καταχωρημένες σε ένα μέσο φιλικό προς τον χρήστη, με πρόσβαση σε όσους ενδιαφέρονται για την ελληνική μετανάστευση στον Καναδά. Τόσο το ΨΑ όσο και η ηλεκτρονική ΒΔ *σχεδιάστηκαν και υλοποιήθηκαν από την ομάδα του Εργαστηρίου Νεοελληνικών Διαλέκτων* στο Πανεπιστήμιο Πατρών, στο πλαίσιο του εν λόγω ερευνητικού προγράμματος.

¹ Διαθέσιμο στο <http://www.meertens.knaw.nl/projecten/sand/sandeng.html/>.

² Διαθέσιμο στο <https://www.dialectsarchive.com/>.

³ Διαθέσιμο στο <https://www.ualberta.ca/~johnnewm/NZEnglish/origins.html/>.

⁴ Διαθέσιμο στο <http://www2.anglistik.uni-freiburg.de/institut/Iskortmann/FRED/>.

⁵ Διαθέσιμο στο <http://research.ncl.ac.uk/necte/>.

⁶ Διαθέσιμο στο <http://www.scottishcorpus.ac.uk/>.

⁷ Διαθέσιμο στο <http://www.tekstlab.uio.no/nota/scandiasyn/>.

⁸ Διαθέσιμο στο <http://www.tgdp.org/>.

⁹ Διαθέσιμο στο <http://www.dialektsyntax.uzh.ch/de.html/>.

¹⁰ Διαθέσιμο στο <http://www.lmgd.philology.upatras.gr/>.

¹¹ Διαθέσιμο στο <http://amigredb.philology.upatras.gr/>.

Η εργασία διαρθρώνεται ως εξής: στην ενότητα μετά την εισαγωγή (Ενότητα 2), παρουσιάζονται τα στάδια συγκρότησης και αρχειοθέτησης του ψηφιακού σώματος τεκμηρίων. Στην επόμενη ενότητα (Ενότητα 3), ακολουθεί η παρουσίαση των τεχνικών προδιαγραφών για την ηλεκτρονική ΒΔ και το ΨΑ του *ImmiGrec*. Η εργασία ολοκληρώνεται με τον επίλογο και στο τέλος παρατίθεται η σχετική βιβλιογραφία.

2. Το ψηφιακό σώμα τεκμηρίων του *ImmiGrec*

Στο πλαίσιο του προγράμματος δημιουργήθηκε ένα ψηφιακό σώμα τεκμηρίων που είναι αποθηκευμένο σε ένα ΨΑ και οργανώνεται σε μια ηλεκτρονική ΒΔ, έτσι ώστε να διευκολύνεται η πρόσβαση στα δεδομένα. Η επεξεργασία των δεδομένων έγινε σύμφωνα με τις επικρατούσες προσεγγίσεις της διαλεκτολογικής –εθνογραφικής– έρευνας (βλ., ενδεικτικά, Hymes 1962, 1964) και της γλωσσολογίας των σωμάτων κειμένων (βλ., για παράδειγμα, το *Μοντέλο 3Α* των Wallis & Nelson 2001).

2.1 Το ψηφιακό σώμα προφορικών κειμένων

2.1.1 Συλλογή προφορικού υλικού

Η συλλογή των προφορικών δεδομένων αποτελεί μία από τις μεγαλύτερες δυσκολίες στη διαλεκτολογική έρευνα, ειδικά όταν οι ερευνητές δεν είναι μέλη της υπό μελέτη γλωσσικής κοινότητας και δεν έχουν κοινωνικούς δεσμούς με αυτή. Προκειμένου να εξαλειφθεί κάθε (επικοινωνιακό) εμπόδιο, η έρευνα πεδίου και η συλλογή δεδομένων πρέπει να βασίζονται στις αρχές της *Εθνογραφίας της Επικοινωνίας* (Ethnography of Communication: Hymes 1962, 1964), όπου η γλώσσα δεν προσλαμβάνεται ως ένας αφηρημένος κώδικας, αλλά ως ένα επιτελεστικό γεγονός που προϋποθέτει ότι μπορεί να αναλυθεί ως ένα σύστημα πρακτικών που καθοδηγείται από κανόνες. Κατά τη συλλογή υλικού, ο ερευνητής πρέπει να λαμβάνει υπόψη τις σημαντικές διαπολιτισμικές διαφορές και δεν πρέπει να έχει προσδοκίες που υπερβαίνουν τις σημασιακές δομές που μοιράζεται η κοινότητα (Stewart & Philipsen 1984).

Υπό αυτό το πρίσμα, η συλλογή των δεδομένων του προγράμματος πραγματοποιήθηκε κυρίως από τους ερευνητές τριών канаδικών πανεπιστημίων που συμμετείχαν στο πρόγραμμα, δηλαδή του McGill, του Simon Fraser και του York, και εν μέρει από τους ερευνητές του Πανεπιστημίου Πατρών, με βάση τη δημιουργία πραγματικών κοινωνικών δεσμών μεταξύ των ερευνητών πεδίου και των γλωσσικών κοινοτήτων. Έγινε καταγραφή της αυθόρμητης και καθημερινής συνομιλίας που υποκινήθηκε μέσω ημιδομημένων συνεντεύξεων, με ιδιαίτερη έμφαση στον φυσικό λόγο, καθώς και άλλα επικοινωνιακά γεγονότα, συμπεριλαμβανομένων των ιστορικών αφηγήσεων. Επιπλέον, οι ερευνητές πεδίου συνέλεξαν κοινωνιογλωσσικές και κοινωνικοϊστορικές πληροφορίες για την καθημερινότητα των μελών των ελληνοκαναδικών κοινοτήτων, ενώ συχνά βασίστηκαν στη βοήθεια μελών των κοινοτήτων.

Για την καθοδήγηση των συνεντεύξεων δημιουργήθηκε ένα ερωτηματολόγιο κοινωνικοϊστορικών και κοινωνιογλωσσικών ζητημάτων, που διαρθρώθηκε γύρω από τέσσερις πυλώνες: (α) την ταυτότητα (αναγνωριστικό ID) του ερευνητή και του πληροφορητή, (β) τον τόπο καταγωγής και αναχώρησης του τελευταίου, (γ) τον τόπο άφιξης και πρώτης εγκατάστασης και (δ) την ενσωμάτωση στην канаδική κοινωνία. Συγκεκριμένα, οι τομείς (β), (γ) και (δ) καλύπτουν τα ακόλουθα θέματα:

- (1) *Καταγωγή και Αναχώρηση*
 - α. Έτος γέννησης
 - β. Τόπος καταγωγής/διαμονής
 - γ. Πόλη/Οικισμός (συγκεκριμένα)
 - δ. Μορφωτικό επίπεδο πριν την αναχώρηση
 - ε. Χρήση γλωσσικής ποικιλίας πριν την αναχώρηση
 - στ. Έτος μετανάστευσης
 - ζ. Μεταφορικό μέσο μετανάστευσης
 - η. Συνολική διάρκεια ταξιδιού (door-to-door)
 - θ. Λόγοι μετανάστευσης
 - ι. Λόγος επιλογής συγκεκριμένης πόλης στο Καναδά
 - ια. Γνώση του Καναδά/Αμερικής πριν την αναχώρηση
 - ιβ. Γνώση Αγγλικής πριν την αναχώρηση
 - ιγ. Γνώση Γαλλικής πριν την αναχώρηση

- (2) *Άφιξη και Εγκατάσταση*
 - α. Πρώτος τόπος εγκατάστασης
 - β. Πόλη/Οικισμός (συγκεκριμένα)
 - γ. Είδος πρώτης απασχόλησης
 - δ. Επαφή με άλλους μετανάστες
 - ε. Επαφή με συντοπίτες
 - στ. Επαφή με (ημι)επίσημους ελληνικούς θεσμούς (προξενείο, κοινότητα, εκκλησία)
 - ζ. Δυσκολία κατά την εγκατάσταση ή/και κατά τις πρώτες επαφές με τις καναδικές αρχές
 - η. Δυσκολία κατά την εγκατάσταση ή/και κατά τις πρώτες επαφές με τη καναδική κοινωνία

- (3) *Ενσωμάτωση*
 - α. Τρέχων τόπος εγκατάστασης
 - β. Πόλη/Οικισμός (συγκεκριμένα)
 - γ. Στάση καναδικής κοινωνίας
 - δ. Σχέση με συναδέλφους
 - ε. Σχέση με άλλες μεταναστευτικές ομάδες
 - στ. Τρέχων μορφωτικό επίπεδο
 - ζ. Καταγωγή συζύγου
 - η. Συμμετοχή σε δραστηριότητες της ελληνικής κοινότητας
 - θ. Συμμετοχή σε δραστηριότητες της εκκλησίας/ενορίας
 - ι. Γλώσσα εργασιακού περιβάλλοντος
 - ια. Γλώσσα οικογενειακού περιβάλλοντος
 - ιβ. Γλώσσα κοινωνικού περιβάλλοντος
 - ιγ. Φοίτηση σε καναδικό σχολείο για την εκμάθηση της Αγγλικής/Γαλλικής
 - ιδ. Συνεισφορά εκμάθησης Αγγλικής/Γαλλικής στην ενσωμάτωση
 - ιε. Γνώση Ελληνικής ή/και διαλέκτου από παιδιά και εγγόνια
 - ιστ. Μορφή χρήσης Ελληνικής ή/και διαλέκτου από παιδιά και εγγόνια
 - ιζ. Παρακολούθηση ελληνικού σχολείου από παιδιά και εγγόνια
 - ιη. Σημασία διατήρησης της ελληνικής γλώσσας
 - ιθ. Διαφορά στη χρήση της Ελληνικής πριν και μετά τη μετανάστευση
 - κ. Αυτοπροσδιορισμός ως Έλληνας/Ελληνίδα ή Καναδός/Καναδή

Δεδομένων των παραπάνω, αλλά και των στόχων του προγράμματος, οι πληροφορητές επιλέχθηκαν σύμφωνα με τα ακόλουθα κοινωνιολογιστικά χαρακτηριστικά:

- (i) την ηλικία (επιλέχθηκαν μέλη των κοινοτήτων 45 ετών και άνω),
- (ii) το φύλο (έγινε προσπάθεια για ισότιμη κατανομή μεταξύ των δύο φύλων),
- (iii) την εκπαίδευση και τον βαθμό γραμματισμού, και
- (iv) τον τόπο καταγωγής στην Ελλάδα.

Από την έρευνα πεδίου συγκεντρώθηκαν περίπου 350 ώρες ηχογραφήσεων και βιντεοσκοπήσεων, εκ των οποίων έχει μεταγραφεί και επισημειωθεί σχεδόν το σύνολο του υλικού. Αξίζει να σημειωθεί ότι στο υλικό μαρτυρούνται σχεδόν όλες οι γλωσσικές ποικιλίες της Ελληνικής (βλ. Πίνακα 1).

Βόρειες διάλεκτοι	Ηπειρωτική
	Θεσσαλική
	Θρακική
	Μακεδονική
	Ρουμελιώτικη
Δωδεκανησιακή	Λεσβιακή
	Σύμη
	Τήλος
	Κάρπαθος
	Κάσος
	Ρόδος
Επτανησιακή	
Κρητική	
Κυκλαδίτικη	
Κυπριακή	
Πελοποννησιακή	
Ποντιακή	
Χιώτικη	

Πίνακας 1: Γλωσσικές ποικιλίες που μαρτυρούνται στο προφορικό υλικό του σώματος τεκμηρίων του *ImmiGrec*

2.1.2 Επισημείωση προφορικού υλικού

Μετά τη συλλογή του πρωτογενούς υλικού ακολούθησε η επισημείωσή του στη ΒΔ και η καταχώρησή του στο ΨΑ του προγράμματος (για τις τεχνικές λεπτομέρειες βλ. Ενότητα 3). Η διαδικασία της επισημείωσης αφορά την καταγραφή συγκεκριμένων μεταδεδομένων σχετικά με τα αρχεία πολυμέσων (συνεντεύξεων), τους πληροφορητές, τις σχετιζόμενες περιοχές, τις γλωσσικές ποικιλίες, τους ερευνητές πεδίου και τις πηγές. Η επισημείωση των μεταδεδομένων έγινε, κυρίως, με βάση τη *δομημένη μέθοδο* (Παπαζαχαρίου & Καρασίμος 2015), η οποία βασίζεται σε προδιαγεγραμμένες κατηγορίες και δομές, καθώς και σε προαποφασισμένα σύνολα τιμών. Με αυτόν τον τρόπο, περιορίζεται ο λανθασμένος προσδιορισμός μεταδεδομένων, καθώς και η λανθασμένη εισαγωγή τιμών τους. Ωστόσο, δεν αποκλείστηκε η δυνατότητα πρόσθεσης νέων τιμών, όταν αυτές δεν είχαν προβλεφθεί εξ αρχής, καθώς και παρουσία πεδίων που απαιτούσαν τη συμπλήρωση κειμένου.

Στον Πίνακα 2 αποτυπώνονται οι πληροφορίες που επισημειώθηκαν για τα αρχεία των συνεντεύξεων στη ΒΔ.

Μεταδεδομένα αρχείων πολυμέσων (συνεντεύξεων)		
Τίτλος αρχείου*		[συμπλήρωση κειμένου]
Πληροφορητής	Κύριος πληροφορητής	[επιλογή τιμής από διευρυνόμενη λίστα]
	Δευτερεύοντες πληροφορητές	[επιλογή τιμής από διευρυνόμενη λίστα]
Ερευνητής Πεδίου*		[επιλογή τιμής από διευρυνόμενη λίστα]
Στοιχεία καταχώρησης*	Αριθμός αρχείων	[συμπλήρωση κειμένου]
	Τύπος αρχείων	[επιλογή τιμής από λίστα]
Μεταγραφή	Μεταγραφέας	[επιλογή τιμής από διευρυνόμενη λίστα]
	Μεταφραστής (στην Αγγλική)	
	Μεταφραστής (στη Γαλλική)	
Ημερομηνία αναφοράς*		[συμπλήρωση κειμένου]
Περιεχόμενα*		[επιλογή τιμής/-ών από (διευρυνόμενη) λίστα]
Σχετιζόμενες περιοχές*		[επιλογή τιμής/ών από διευρυνόμενη λίστα]
Άδειες*	Άδειες χρήσης	[επιλογή τιμής από λίστα]
	Πηγή	[επιλογή τιμής από διευρυνόμενη λίστα]
Λέξεις-κλειδιά		[συμπλήρωση κειμένου]

Πίνακας 2: Επισημείωση μεταδεδομένων για τα αρχεία των συνεντεύξεων

Στον Πίνακα 3 αποτυπώνονται οι πληροφορίες που επισημειώθηκαν για τους πληροφορητές στη ΒΔ. Οι πληροφορίες αυτές αποδελτιώνουν τις απαντήσεις που προέκυψαν από το ερωτηματολόγιο που συντάχθηκε για τη διεξαγωγή των συνεντεύξεων (βλ. 1-3).

Μεταδεδομένα πληροφορητών		
Προσωπικές πληροφορίες*	Όνομα*	[συμπλήρωση κειμένου]
	Επώνυμο*	
	Έτος γέννησης*	
	Φύλο	
Καταγωγή	Τόπος καταγωγής	[επιλογή τιμής από διευρυνόμενη λίστα]
	Γλωσσική ποικιλία	
	Βαθμός χρήσης γλωσσικής ποικιλίας	[επιλογή τιμής από λίστα]
	Μορφωτικό επίπεδο	
Αναχώρηση	Λόγοι μετανάστευσης	[επιλογή τιμής/-ών από λίστα]
	Λόγοι επιλογής πόλης στον Καναδά	
	Βαθμός γνώσης του Καναδά	
	Βαθμός γνώσης της Αγγλικής	
	Βαθμός γνώσης της Γαλλικής	
Ταξίδι	Έτος μετανάστευσης	[συμπλήρωση κειμένου]
	Μέσο μεταφοράς	[επιλογή τιμής/-ών από λίστα]
	Συνολικός χρόνος ταξιδιού	[συμπλήρωση κειμένου]
	Αφήγηση ταξιδιού μετανάστευσης	[συμπλήρωση κειμένου]

Αφιξη/ Εγκατάσταση	Πρώτος τόπος εγκατάστασης	[επιλογή τιμής από διευρυνόμενη λίστα]
	Είδος πρώτης εργασίας	[επιλογή τιμής/-ών από λίστα]
	Βαθμός επαφής με άλλους μετανάστες	[επιλογή τιμής από λίστα]
	Βαθμός επαφής με Έλληνες μετανάστες	
	Βαθμός επαφής με (ημι)επίσημα ελληνικά ιδρύματα (π.χ. προξενείο, κοινότητα, εκκλησία)	
	Βαθμός δυσκολίας κατά την εγκατάσταση ή/και κατά τη διάρκεια της πρώτης επαφής με τις καναδικές αρχές	
	Βαθμός δυσκολίας κατά την εγκατάσταση ή/και κατά τη διάρκεια της πρώτης επαφής με την καναδική κοινωνία	
	Αφήγηση κάποιου αξιοσημείωτου περιστατικού κατά την άφιξη/εγκατάσταση	[συμπλήρωση κειμένου]
Κοινωνική ενσωμάτωση	Τρέχων τόπος κατοικίας	[επιλογή τιμής από διευρυνόμενη λίστα]
	Στάση καναδικής κοινωνίας	[επιλογή τιμής από λίστα]
	Βαθμός επαφής με τους συναδέλφους	
	Βαθμός επαφής με άλλους μετανάστες	
	Τρέχων μορφωτικό επίπεδο	
	Καταγωγή συζύγου	
	Βαθμός συμμετοχής σε δραστηριότητες της ελληνικής κοινότητας	
	Βαθμός συμμετοχής σε δραστηριότητες της εκκλησίας/ενορίας	
	Συμμετοχή σε κινήματα κατά της δικτατορίας (ναι/όχι)	
	Αυτοπροσδιορισμός ως Έλληνας/-ίδα ή Καναδός/-ή	
Γλωσσική ενσωμάτωση	Γλώσσα στον χώρο εργασίας	[επιλογή τιμής/-ών από λίστα]
	Γλώσσα στο οικογενειακό περιβάλλον	
	Γλώσσα στο κοινωνικό περιβάλλον	
	Φοίτηση σε καναδικό σχολείο για την εκμάθηση αγγλικών/γαλλικών	[επιλογή τιμής από λίστα]
	Αφήγηση ενός γεγονότος σχετικά με τη χρήση γλώσσας	[συμπλήρωση κειμένου]
	Βαθμός συμβολής της εκμάθησης Αγγλικών/Γαλλικών στην κοινωνική ενσωμάτωση	[επιλογή τιμής από λίστα]
	Βαθμός γνώσης Ελληνικής και/ή διαλέκτου από παιδιά και εγγόνια	
	Μορφή χρήσης της Ελληνικής και/ή της διαλέκτου από παιδιά και εγγόνια	
	Φοίτηση παιδιών και εγγονών σε ελληνικά σχολεία (ναι/όχι)	
	Βαθμός σημασίας της διατήρησης της Ελληνικής	
Βαθμός διαφοράς χρήσης της Ελληνικής πριν και μετά τη μετανάστευση		
Αφήγηση ιστορίας στη διαλεκτική ποικιλία	[συμπλήρωση κειμένου]	

Πίνακας 3: Επισημείωση μεταδεδομένων για τους πληροφορητές

Τέλος, οι οντότητες της ΒΔ που αντιστοιχούν σε διευρυνόμενες λίστες τιμών απεικονίζονται στον Πίνακα 4.

Σχετιζόμενες περιοχές	Όνομα*	[συμπλήρωση κειμένου]
	Περιγραφή	
Γλωσσικές ποικιλίες	Όνομα*	[συμπλήρωση κειμένου]
	Περιγραφή	
Ερευνητές πεδίου	Πανεπιστήμιο*	[επιλογή τιμής από λίστα]
	Προσφώνηση*	
	Όνομα*	[συμπλήρωση κειμένου]
	Επίθετο*	
Πηγές	Όνομα*	[συμπλήρωση κειμένου]
	Περιγραφή	

Πίνακας 4: Μεταδεδομένα για τις οντότητες που αντιστοιχούν στις διευρυνόμενες λίστες τιμών

Τα πεδία στους Πίνακες 2, 3, 4 που σημειώνονται με αστερίσκο είναι υποχρεωτικά. Η συμπλήρωση κάθε πεδίου γίνεται είτε με την επιλογή από μια (διευρυνόμενη) λίστα τιμών είτε με τη συμπλήρωση κειμένου. Στην περίπτωση των διευρυνόμενων λιστών (βλ. λίστες με περιοχές, γλωσσικές ποικιλίες, ερευνητές και πηγές), η προσθήκη νέων τιμών γίνεται ελεύθερα από τους ίδιους τους χρήστες του συστήματος. Αντίθετα, η προσθήκη νέων τιμών στη διευρυνόμενη λίστα του πεδίου «Περιεχόμενα» (βλ. Πίνακα 2) μπορεί να γίνει μόνο μετά από επικοινωνία με τον διαχειριστή του συστήματος.

Η επισημείωση αποτελεί ουσιαστική συνεισφορά στη μελέτη ενός τόσο εκτεταμένου σώματος τεκμηρίων. Πρόκειται για ένα εργαλείο μοναδικό για τη γλωσσική, ιστορική ή άλλου είδους έρευνα, η οποία μπορεί να πραγματοποιείται ηλεκτρονικά και με βάση τις επισημειωμένες κατηγορίες (βλ., για παράδειγμα, την εξαγωγή στατιστικών στη ΒΔ, Ενότητα 3.1). Φιλοδοξία των συνεργατών του προγράμματος είναι η συγκεκριμένη ΒΔ να αποτελέσει πολύτιμο εργαλείο έρευνας σε κάθε επίπεδο.

2.1.3 Μεταγραφή προφορικού υλικού

Μετά τη συλλογή του πρωτογενούς προφορικού υλικού, την οργάνωσή του στη ΒΔ και την καταχώρησή του στο ΨΑ, το επόμενο στάδιο της επεξεργασίας αφορούσε τη μεταγραφή του (σχετίζεται με τη γραπτή απόδοση των προφορικών συνεντεύξεων), ώστε να διευκολυνθεί η μελέτη και επεξεργασία του υλικού. Συγκεκριμένα, έχει γίνει μεταγραφή του συνόλου σχεδόν του υλικού που έχει συγκεντρωθεί μέσω των συνεντεύξεων. Τα αρχεία των συνεντεύξεων μεταγράφηκαν «δια χειρός», χωρίς τη βοήθεια αυτοματοποιημένου λογισμικού μεταγραφής (*μετατροπή φωνής σε κείμενο* 'voice-to-text'), λόγω των δυσκολιών που ενδεχομένως να προκαλούσε ένα τέτοιο εγχείρημα (π.χ. δυσκολία αυτόματης αναγνώρισης διαλεκτικού λόγου, δυσκολία αυτόματης αναγνώρισης εναλλαγής κωδικών και δάνειων λέξεων, καθώς και της ομιλίας υπερήλικων ατόμων κ.ά.).

Στις μέρες μας, το είδος της μεταγραφής που έχει προκριθεί σε πολλές βάσεις δεδομένων είναι η ορθογραφική (βλ. Anderwald & Wagner 2007· Παπαζαχαρίου & Καρασίμος 2015). Μολονότι είναι η λιγότερο πιστή, σε σχέση με την απόδοση των ήχων που χρησιμοποιούν πραγματικά οι ομιλητές, οι μεταγραφείς δεν χρειάζονται περαιτέρω εκπαίδευση και εξειδίκευση, αποτελεί τον πιο γρήγορο τύπο μεταγραφής σε σχέση με όλους τους άλλους και, τέλος, είναι πιο κατανοητή στο ευρύ κοινό. Παρόλα αυτά, εάν

κάποιοι ενδιαφέρεται για ειδικά γλωσσικά φαινόμενα, η παράλληλη εμφάνιση των πρωτογενών δεδομένων με την ορθογραφική μεταγραφή τους επιτρέπει τη χρήση της τελευταίας ως οδηγού για τον εντοπισμό των φαινομένων αυτών στα πρωτογενή δεδομένα, καθώς και τον έλεγχο της ακριβούς πραγμάτωσής τους.

Η επιλογή της ορθογραφικής μεταγραφής συνοδεύτηκε από κάποιες συμβάσεις λόγω των ιδιότυπων χαρακτηριστικών της γλώσσας πολύγλωσσων μεταναστών. Πιο συγκεκριμένα, στην περίπτωση του προφορικού διαλεκτικού λόγου, χρησιμοποιήθηκε η ορθογραφική μεταγραφή, συνοδευόμενη από ειδικά σύμβολα που υποδεικνύουν διαλεκτικά φαινόμενα, κυρίως φωνολογικά, που απαντώνται στα όρια των λέξεων. Για παράδειγμα:

(4) *Εχ' αβάντα του-μπιθιρό-τ* 'Στηρίζεται στον πεθερό του' (Λεσβιακά)

Όσον αφορά την *εναλλαγή κωδίκων* (code-switching) μεταξύ ελληνικών, αγγλικών και γαλλικών, που ήταν αναμενόμενη στην περίπτωση των πολύγλωσσων μεταναστών, οι μεταγραφείς ακολούθησαν και πάλι τον δρόμο της ορθογραφικής μεταγραφής στην αντίστοιχη γλώσσα (5α και 5β). Ωστόσο, όταν οι ξενικές λέξεις παρουσίαζαν στοιχεία ενσωμάτωσης στη γλώσσα-δέκτη (δηλαδή την Ελληνική), η ορθογραφική μεταγραφή γινόταν με βάση την ελληνική ορθογραφία (5γ). Για παράδειγμα:

- (5) α. Λέει «Πότε... πότε γεννήθηκες;». Λέω «Το σαράντα έξι». Ω, λέει «Είσαι *thirteen*». Λέω «*No*, είμαι *fourteen*».
 β. If it's business-wise, I will speak English. If it's a social-wise, *parlez vous français*.
 γ. (...) τότε με πήρε με πήγε στον σταθμό, μου 'κοψε το *τικέτο*.

Η μεταγραφή του προφορικού υλικού σε γραπτό κείμενο αποτελεί σημαντική συνεισφορά που, σε συνδυασμό με τα επισημειωμένα μεταδεδομένα, διευκολύνει τη μελέτη του συγκεντρωμένου υλικού τόσο για τη στοχευμένη (γλωσσολογική, ιστορική, εθνογραφική κ.ά.) έρευνα όσο και για αυτούς που ενδιαφέρονται ευρύτερα για την ελληνική μετανάστευση στον Καναδά.

2.2 Άλλοι τύποι τεκμηρίων

Το ΨΑ και η ΒΔ του προγράμματος εκτός από γλωσσικό υλικό διαθέτουν, επίσης, πλούσιο αρχειακό υλικό που αφορά σε διάφορες εκφάνσεις της μετανάστευσης των Ελλήνων στον Καναδά. Πιο συγκεκριμένα, έχουν καταχωρηθεί έως τώρα 439 αρχεία εικόνας και 10 αρχεία βίντεο. Αυτά τα αρχεία αφορούν στις εξής κατηγορίες:

- (6) *Περιεχόμενα αρχείων πολυμέσων*
- | | |
|-----------------------------------|-------------------------------------|
| α. Έντυπα αιτήσεων | ια. Συνεντεύξεις |
| β. Δράση κατά της δικτατορίας | ιβ. Διαλέξεις/ομιλίες |
| γ. Ημερολόγια κοινοτικών συλλόγων | ιγ. Επιστολές |
| δ. Ταχυδρομικές κάρτες | ιδ. Αποκόμματα αποσκευών |
| ε. Εορτασμοί | ιε. Χάρτες |
| στ. Τηλεοπτικό/ραδιοφωνικό υλικό | ιστ. Ιατρικές εξετάσεις |
| ζ. Οικογενειακές φωτογραφίες | ιζ. Εφημερίδες/περιοδικά |
| η. Φωτογραφίες προσώπων | ιη. Παλαιότερες/σύγχρονες αφηγήσεις |
| θ. Εθνικές εορτές | ιθ. Διαβατήρια |
| ι. Κάρτες διακοπών | ικ. Προσωπικά ημερολόγια |

κα. Αφίσες πλοίων με προορισμό τον Καναδά	κγ. Θρησκευτικές εκδηλώσεις
κβ. Σχέσεις με άλλες εθνοτικές ομάδες	κδ. Εισιτήρια
	κε. Ξενοφοβική συμπεριφορά

Όπως έχει ήδη αναφερθεί, η παραπάνω λίστα δεν είναι εξαντλητική, αλλά μπορεί να διευρύνεται με την προσθήκη νέων τιμών για διαφορετικά περιεχόμενα. Η επισημείωση των υπόλοιπων πληροφοριών για αυτού του τύπου τα τεκμήρια αντιστοιχεί σε αυτή για τα αρχεία των συνεντεύξεων (βλ. Πίνακα 2).

3. Η ΒΔ και το ΨΑ του *ImmiGrec*

Όλα τα τεκμήρια βρίσκονται σε δύο διαφορετικές διαδικτυακές υπηρεσίες, οι οποίες έχουν στηθεί αλληλένδετα και εξυπηρετούν πλήρως την καταχώρηση όλων των αρχείων και των αντίστοιχων μεταδεδομένων τους. Τα αρχεία καταχωρούνται στο ΨΑ, ενώ τα μεταδεδομένα στη συνοδευτική ΒΔ. Και οι δύο υπηρεσίες παρέχονται μέσω της ίδιας τεχνολογικής υποδομής και γίνεται χρήση της ευρέως γνωστής στοίβας εφαρμογών *Linux/Apache/MySQL/PHP* (γνωστή ως LAMP, βλ. Gerner *et al.* 2005).

3.1 Η ΒΔ του *ImmiGrec*

Η ΒΔ έχει στηθεί σε μια ειδικά κατασκευασμένη πλατφόρμα με το Drupal 7 CMS (v.7), ένα σύστημα διαχείρισης υλικού ανοιχτού κώδικα (open source content management system) το οποίο υποστηρίζεται ενεργά από χιλιάδες προγραμματιστές, παρέχοντας συνεχώς υψηλή ασφάλεια ενάντια σε κακόβουλες ηλεκτρονικές επιθέσεις. Οι οντότητες του συστήματος είναι οι εξής:

- (i) Αρχεία πολυμέσων (βλ. Πίνακα 2)
- (ii) Πληροφορητές (βλ Πίνακα 3)
- (iii) Σχετιζόμενες περιοχές (βλ. Πίνακα 4)
- (iv) Γλωσσικές ποικιλίες (βλ. Πίνακα 1 & 4)
- (v) Ερευνητές πεδίου (βλ. Πίνακα 4)
- (vi) Πηγές (βλ. Πίνακα 4)

Για κάθε εργασία ανάλυσης των αρχείων ήχου και βίντεο γίνεται χρήση της εφαρμογής *Ffmpeg*¹² μέσω της γραμμής εντολών, η οποία βοηθά στην εξαγωγή ποιοτικών και ποσοτικών χαρακτηριστικών. Συγκεκριμένα, παρέχει την κωδικοποίηση του αρχείου, καθώς επίσης και τη διάρκεια κάθε αρχείου σε δευτερόλεπτα.

Η ΒΔ συνοδεύεται από επιπλέον σελίδες ειδικού σκοπού, όπως:

- (i) Σελίδα προβολής αρχείου
- (ii) Σελίδα σφαλμάτων
- (iii) Σελίδα ελλিপών αρχείων
- (iv) Σελίδα προβολής στατιστικών πληροφορητών
- (v) Σελίδα στατιστικών ανά ερευνητή και ίδρυμα
- (vi) Σελίδα προβολής καταχωρημένων συνεντεύξεων

¹² Διαθέσιμο στο <https://www.ffmpeg.org/>.

Immigration and Language in Canada

Media files Informants Places Linguistic varieties Field researchers Sources

Home » Media files » interview (808)

View Edit Revisions Devel

Submitted by Siotou Alexandra on Thu, 30/08/2018 - 22:12

Field researcher: Siotou Alexandra, Dr.
Media file ID: 808
Parts: 1
Main informant: Iladi Eleni
Type of file: Audio (.flac)
Contents: Interview
Date of reference: Wednesday, 1 January, 1958
Permitted use: Free for public use
Folder: 2018/08/000808
Places of interest: Canada, Montreal, Athens
Source: Immigration and Language in Canada, project

File name guidelines

This is a list of files that should *or could* be placed under the folder `/ / rawdata / 2018 / 08 / 000808 / at cloud.immigrec.com`.
Caution. file names are case sensitive, meaning `.wav` is different than `.WAV` or `.Wew`.

- **View** `000808-439.flac` This is the name of the main file, before you upload it.
- `000808-439.txt` Various details in plain text
- `000808-439.txt@1d` Textgrid file
- Language: English
 - `000808-439-en.pdf` Transcript in PDF format
 - `000808-439-en.tex` Transcript in TEX format
- Language: French
 - `000808-439-fr.pdf` Transcript in PDF format
 - `000808-439-fr.tex` Transcript in TEX format
- Language: Greek
 - `000808-439-el.pdf` Transcript in PDF format
 - `000808-439-el.tex` Transcript in TEX format

Uploaded files

000808-439.flac

[Download raw file](#)

Input #0, flac, from /2018/08/000808/000808-439.flac:
 Duration: 01:19:23.27, start: 0.000000, bitrate: 725 kb/s
 Stream #0:0 Audio: flac, 44100 Hz, stereo, s16

▶ 0:00 / 0:00

Add new comment

Comments

Submitted by Chairetakis George on Fri, 28/09/2018 - 13:42 Permalink

Media file changed from WAV

Media file changed from WAV to FLAC.
 Noise reduction process has been effectively applied.

delete edit reply

Add new comment

Your name Tsimpouris Char...

Comment *

[More information about text formats](#)

Text format Filtered

- Web page addresses and e-mail addresses turn into links automatically.
- Allowed HTML tags: `<a>` `` `` `<code>` `` `<u>` `<sup>` `<sub>` `<hr>` `<div>` `<div>`
- Lines and paragraphs break automatically.

In case you want to start a discussion for this media file, or reply to an older post, fill in your answer on top and then click Save.

Εικόνα 1: Σελίδα προβολής καταχωρημένου τεκμηρίου

Κάθε καταχωρημένο αρχείο συνοδεύεται από μία σελίδα προβολής η οποία βοηθά τον χρήστη, αρχικά, να ανεβάσει το αρχείο και, στη συνέχεια, να το προβάλει εύκολα μέσα από τον φυλλομετρητή, χωρίς την ανάγκη χρήσης του ΨΑ. Συγκεκριμένα, στην *Εικόνα 1* απεικονίζονται τα επιμέρους στοιχεία που ακολουθούν:

- (i) **A:** Πληροφορίες σχετικά με το αρχείο, τον χρήστη που έκανε την καταχώρηση, τη σύνδεση με τον πληροφορητή
- (ii) **B:** Οδηγίες για τη σωστή αποθήκευση του αρχείου στο ΨΑ
- (iii) **Γ:** Γρήγορη προεπισκόπηση του αρχείου (αρχείο ήχου στο παράδειγμα)
- (iv) **Δ:** Σχόλια άλλων χρηστών της ΒΔ για το συγκεκριμένο αρχείο
- (v) **Ε:** Φόρμα προσθήκης νέου σχολίου από τον τρέχοντα χρήστη για το συγκεκριμένο αρχείο

Η σελίδα προβολής σφαλμάτων για τα καταχωρημένα τεκμήρια (Εικόνα 2) εμφανίζει μία λίστα με αρχεία τα οποία βρίσκονται ήδη στο ΨΑ, αλλά η ονομασία τους ή η θέση τους σε φάκελο είναι εσφαλμένη. Το σφάλμα μπορεί να σχετίζεται με τυπογραφικό λάθος ή με εσφαλμένη χρήση του συστήματος, καθιστώντας αδύνατη τη σύνδεση της πληροφορίας με τη ΒΔ. Η σελίδα σφαλμάτων αποτελεί έναν εύκολο τρόπο γρήγορου ελέγχου και επιβεβαίωσης της έγκυρης κατάστασης του ΨΑ ανά πάσα στιγμή. Επίσης, η λίστα περιλαμβάνει και ορισμένες τεχνικές εργασίες αναφορικά με την εξοικονόμηση χώρου αποθήκευσης στο ΨΑ, χωρίς όμως έκπτωση της ποιότητας. Αυτό επιτυγχάνεται με τη μετατροπή αρχείων ήχου από ασυμπίεστη μορφή WAV σε συμπιεσμένη μορφή FLAC.



Immigration and Language in Canada

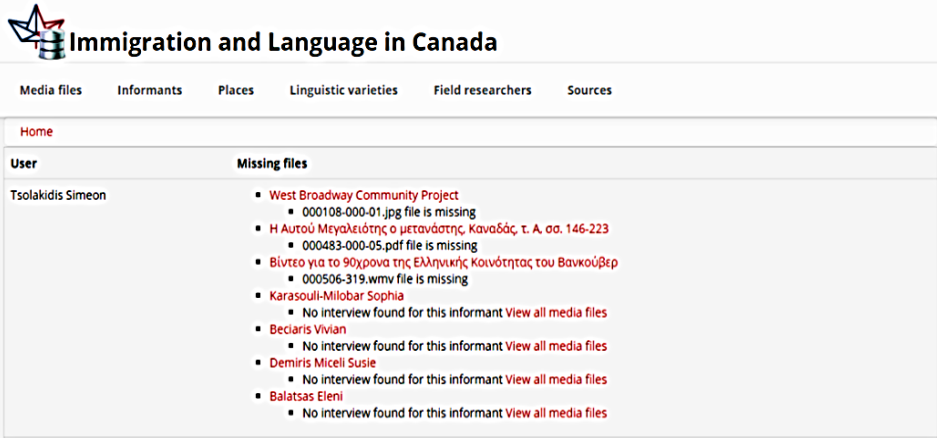
Media files Informants Places Linguistic varieties Field researchers Sources

Home » Error files » Check again

Folder/File under Rawdata	Error
/2018/01/000825/IMG_0025.jpg	<ul style="list-style-type: none"> Unable to process file in order to identify media details. Please correct or contact the person who uploaded the file (shown through cloud web page). File names must use only small letters. Please correct to img_0025.jpg .
/2018/02/000642/000642-356-el.tex.tex	<ul style="list-style-type: none"> Filename is not correct. It should be NAME.EXT
/2018/10/000826/IMG_0029.JPG	<ul style="list-style-type: none"> Unable to process file in order to identify media details. Please correct or contact the person who uploaded the file (shown through cloud web page). File names must use only small letters. Please correct to img_0029.jpg .
/2018/10/000826/IMG_0030.JPG	<ul style="list-style-type: none"> Unable to process file in order to identify media details. Please correct or contact the person who uploaded the file (shown through cloud web page). File names must use only small letters. Please correct to img_0030.jpg .
/2018/10/000827/IMG_0031.JPG	<ul style="list-style-type: none"> Unable to process file in order to identify media details. Please correct or contact the person who uploaded the file (shown through cloud web page). File names must use only small letters. Please correct to img_0031.jpg .

Εικόνα 2: Σελίδα προβολής σφαλμάτων για καταχωρημένα τεκμήρια

Η σελίδα ελλিপών αρχείων προβάλλει αναλυτικά τις περιπτώσεις κατά τις οποίες, ενώ έχουν καταχωρηθεί τα μεταδεδομένα τεκμηρίων στη ΒΔ, οι αντίστοιχες καταχωρήσεις των αρχείων στο ΨΑ εκκρεμούν (βλ. Εικόνα 3).



Immigration and Language in Canada

Media files Informants Places Linguistic varieties Field researchers Sources

Home

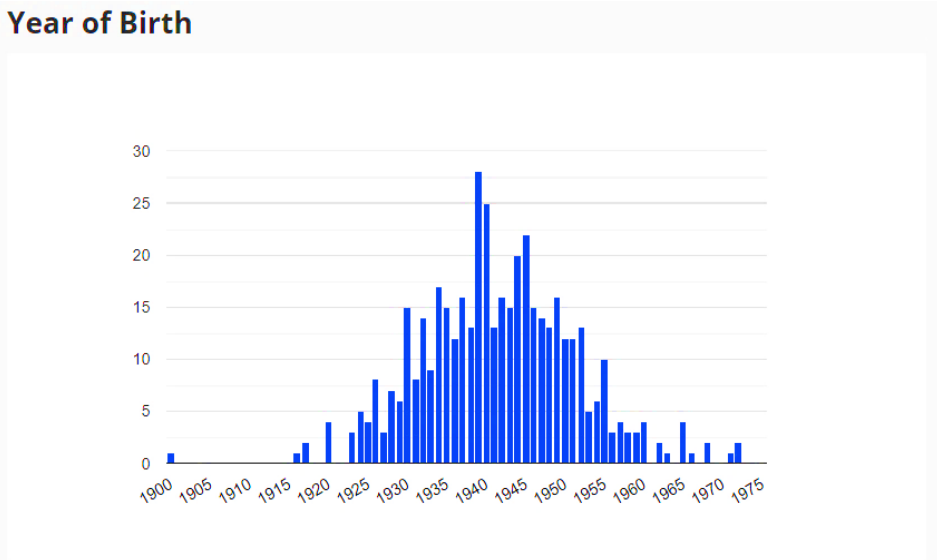
User Missing files

Tsolakidis Simeon

- West Broadway Community Project
 - 000108-000-01.jpg file is missing
- Η Αυτού Μεγαλειότης ο μετανάστης Καναδάς, τ. Α. σσ. 146-223
 - 000483-000-05.pdf file is missing
- Βίντεο για το 90χρονα της Ελληνικής Κοινότητας του Βανκούβερ
 - 000506-319.wmv file is missing
- Karasouli-Milobar Sophia
 - No interview found for this informant [View all media files](#)
- Beclaris Vivian
 - No interview found for this informant [View all media files](#)
- Demiris Miceli Susie
 - No interview found for this informant [View all media files](#)
- Balatsas Eleni
 - No interview found for this informant [View all media files](#)

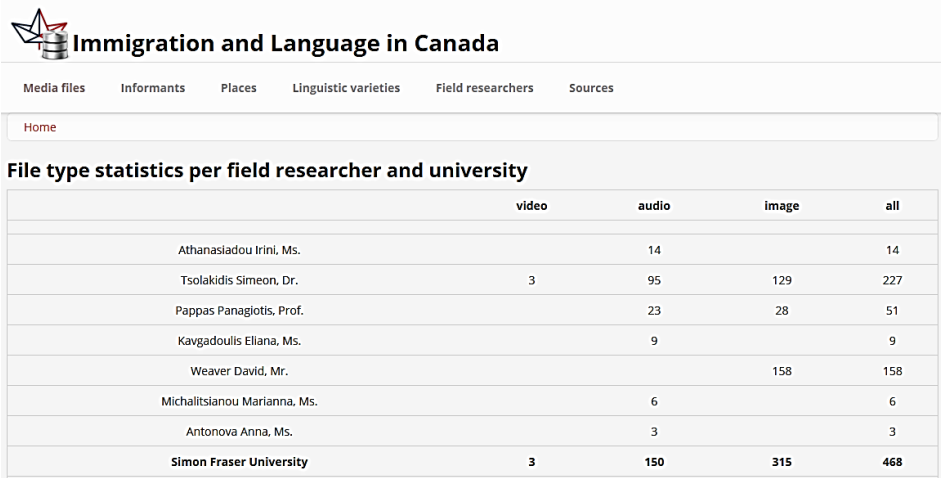
Εικόνα 3: Σελίδα προβολής ελλিপών αρχείων

Η σελίδα προβολής στατιστικών για τους πληροφορητές παρέχει με τη μορφή γραφημάτων όλα τα καταχωρημένα μεταδεδομένα ανά πεδίο για όλους τους πληροφορητές (βλ. Πίνακα 3). Στην *Εικόνα 4* δίνεται ένα ενδεικτικό παράδειγμα για την κατανομή του έτους γέννησης όλων των πληροφορητών της ΒΔ.



Εικόνα 4: Σελίδα προβολής στατιστικών για τους πληροφορητές (κατανομή έτους γέννησης)

Η σελίδα προβολής στατιστικών ανά ερευνητή και ίδρυμα προσφέρει ποσοτικές πληροφορίες σχετικά με τις καταχωρήσεις, καθώς και την εποπτική εικόνα του όγκου των εργασιών ανά ίδρυμα. Στην *Εικόνα 5* δίνεται ένα δείγμα.

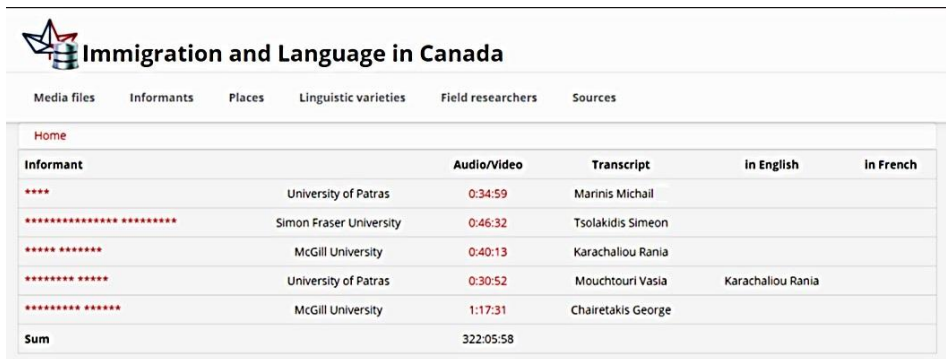


The screenshot shows a website header with the title 'Immigration and Language in Canada' and a navigation menu with items: Media files, Informants, Places, Linguistic varieties, Field researchers, and Sources. Below the header is a 'Home' link. The main content area is titled 'File type statistics per field researcher and university' and contains a table with the following data:

	video	audio	image	all
Athanasiadou Irini, Ms.		14		14
Tsolakidis Simeon, Dr.	3	95	129	227
Pappas Panagiotis, Prof.		23	28	51
Kavgadoulis Eliana, Ms.		9		9
Weaver David, Mr.			158	158
Michalitsianou Marianna, Ms.		6		6
Antonova Anna, Ms.		3		3
Simon Fraser University	3	150	315	468

Εικόνα 5: Σελίδα προβολής στατιστικών ανά ερευνητή και ίδρυμα

Τέλος, στη σελίδα καταχωρημένων συνεντεύξεων, προβάλλονται βασικές πληροφορίες για τη διάρκεια των ηχητικών αρχείων (και βίντεο) των συνεντεύξεων και για το αν έχει πραγματοποιηθεί η αντίστοιχη μεταγραφή (βλ. Εικόνα 6).



The screenshot shows the same website header as above. Below the header is a 'Home' link. The main content area is titled 'Audio/Video' and contains a table with the following data:

Informant	Audio/Video	Transcript	In English	In French
****	University of Patras	0:34:59	Marinis Michail	
*****	Simon Fraser University	0:46:32	Tsolakidis Simeon	
*****	McGill University	0:40:13	Karachaliou Rania	
*****	University of Patras	0:30:52	Mouchtouris Vasia	Karachaliou Rania
*****	McGill University	1:17:31	Chairetakis George	
Sum		322:05:58		

Εικόνα 6: Σελίδα προβολής καταχωρημένων συνεντεύξεων

3.2 Το ΨΑ του *ImmiGrec*

Η ΒΔ συνοδεύεται από το ΨΑ στο οποίο γίνεται η καταχώρηση των αρχείων μετά την επισημείωση των μεταδεδομένων. Για το ΨΑ έγινε χρήση της πλατφόρμας ownCloud¹³ η οποία είναι ανοιχτού κώδικα και ελεύθερα διαθέσιμη στο διαδίκτυο. Η επιλογή της συγκεκριμένης πλατφόρμας συνοψίζεται στους εξής λόγους:

- (i) Βρίσκεται στην έκδοση 9 που σημαίνει πως έχει ωριμάσει αρκετά και αποτελεί μια αξιόπιστη επιλογή.

¹³ Διαθέσιμο στο <https://owncloud.org/>.

- (ii) Υποστηρίζεται από εκατοντάδες προγραμματιστές και χρησιμοποιείται από χιλιάδες χρήστες. Ως εκ τούτου, έχει ελεγχθεί ενδελεχώς για κενά ασφαλείας, ενώ παράλληλα υλοποιούνται πολλές από τις απαιτήσεις πραγματικών χρηστών.
- (iii) Υποστηρίζει κοινόχρηστους φακέλους μεταξύ των χρηστών.
- (iv) Επιτρέπει τη ρύθμιση των κοινόχρηστων αρχείων με δικαιώματα μόνο για ανάγνωση, αν αυτό είναι αναγκαίο.
- (v) Είναι συμβατή με κινητές συσκευές που διαθέτουν λειτουργικό iOS και Android, αλλά και λειτουργικά συστήματα Windows και MacOS.

4. Επίλογος

Το ΨΑ και η ΒΔ του προγράμματος *ImmiGrec* αποτελούν έναν σημαντικό πόρο συγκεντρωμένου και ταξινομημένου υλικού (πάνω από 680 τεκμήρια και 350 ώρες ηχογραφήσεων/βιντεοσκοπήσεων μαζί με τις μεταγραφές και την κωδικοποίηση ενός μεγάλου αριθμού μεταδεδομένων) που καλύπτει το κενό που υπήρχε στη μελέτη της ελληνικής διατλαντικής μετανάστευσης, καθώς και στην κατανόηση της εθνοτικής πολυμορφίας στην канаδική κοινωνία. Με άλλα λόγια, αυτά τα δύο αλληλένδετα ψηφιακά προϊόντα παρέχουν από κοινού το υπόβαθρο και τα εργαλεία για τη διερεύνηση της γλώσσας και της ιστορίας των Ελλήνων μεταναστών στον Καναδά, αλλά και την αποσαφήνιση της σχέσης τους με τη γλωσσική, κοινωνική και πολιτιστική ιστορία της χώρας. Ως εκ τούτου, το έργο προωθεί τη διεπιστημονικότητα μεταξύ της γλωσσολογικής, ιστορικής και κοινωνιολογικής έρευνας, προκειμένου να καλύψει το κενό που υπήρχε, αλλά και να παράσχει τη δυνατότητα μιας πλήρως αναπτυγμένης ανάλυσης τόσο των (κοινωνιο)γλωσσικών όσο και των (κοινωνικο)ιστορικών χαρακτηριστικών των ελληνοκαναδικών κοινοτήτων. Τέλος, η συγκεκριμένη προσπάθεια προσδίδει και μια ιστορική κοινωνιολογική διάσταση στο αντικείμενο του προγράμματος και συντελεί στη διαχρονική μελέτη του.

Είναι αξιοσημείωτο ότι το παρόν έργο επιχειρεί να προχωρήσει ένα βήμα πέρα από την καθιερωμένη πρακτική στη γλωσσική και ιστορική έρευνα, για να διασφαλίσει τη βιωσιμότητα των αποτελεσμάτων, με την παραγωγή πρωτοποριακών ψηφιακών προϊόντων που σχετίζονται με την ανάπτυξη των Ψηφιακών Ανθρωπιστικών Επιστημών. Σε αυτό το πνεύμα, χρησιμοποιήθηκαν νέες τεχνολογίες για τη δημιουργία μιας συνεχώς διευρυνόμενης βάσης δεδομένων, η οποία διευκολύνει την πρόσβαση σε όποιον ενδιαφέρεται να μελετήσει τις διάφορες γλωσσικές, ιστορικές και κοινωνικές πτυχές των ελληνοκαναδικών κοινοτήτων, με βασικά ερευνητικά κριτήρια και κατηγορίες, όπως φύλο, τόπος προέλευσης, ημερομηνία άφιξης, τόπος εγκατάστασης, είδος εργασίας κ.λπ.

Εν κατακλείδι, η δημιουργία του ΨΑ και της ΒΔ του *ImmiGrec* εγείρει την επίγνωση της ιστορίας και της κληρονομιάς των Ελληνοκαναδών και ενισχύει τόσο τις σχέσεις εντός της κοινότητας των μεταναστών όσο και τις διμερείς σχέσεις μεταξύ των δύο χωρών, αυξάνοντας το ενδιαφέρον του ευρύ κοινού, αλλά και παρέχοντας ένα σταθερό σημείο αναφοράς για εκπαιδευτικούς, ακαδημαϊκούς ή άλλους σκοπούς.

Ευχαριστίες

Το έργο υλοποιήθηκε στο πλαίσιο του ερευνητικού προγράμματος *ImmiGrec* (Ιαν. 2017 - Δεκ. 2018), το οποίο χρηματοδοτήθηκε από το *Ίδρυμα Σταύρος Νιάρχος* και απασχόλησε ερευνητικές ομάδες

από τέσσερα πανεπιστήμια σε Ελλάδα και Καναδά (Πανεπιστήμιο Πατρών, McGill University, York University και Simon Fraser University).

Θα θέλαμε να ευχαριστήσουμε θερμά το *Τόρμα Σταύρος Νιάρχος* για την ευγενή του χορηγία και τους συνεργάτες μας στον Καναδά και τις ερευνητικές τους ομάδες, συγκεκριμένα, τους Τάσο Αναστασιάδη (McGill University), Σάκη Γκέκα (York University) και Παναγιώτη Παππά (Simon Fraser University) για τη συνεργασία τους και το υλικό που έχουν συλλέξει στο πλαίσιο του προγράμματος, μέρος του οποίου χρησιμοποιήσαμε στα παραδείγματα της συγκεκριμένης εργασίας.

Βιβλιογραφία

- Anderson, W. & J. Corbett. 2009. The SCOTS corpus: a users' guide. *Scottish language* 27: 19-41.
- Anderwald, L. & S. Wagner. 2007. FRED-the Freiburg English Dialect Corpus. Στο J. Beal, P. Corrigan & H. Moisl (επιμ.), *Creating and digitizing language corpora*, Vol. 1: Synchronic databases. London: Palgrave Macmillan, 35-53.
- Barbiers, S. et al. 2006. *Dynamic Syntactic Atlas of the Dutch dialects (DynaSAND)*. Amsterdam: Meertens Institute.
- Bayard, D. 2000. New Zealand English: origins, relationships, and prospects. *Moderna Språk* 94(1): 8-14.
- Beal, J. C., Corrigan, K. P. & H. L. Moisl (επιμ.). 2007. *Creating and digitizing language corpora*, Vol. 1 & 2. Basingstoke: Palgrave Macmillan.
- Boas, H. C. 2002. The Texas German Dialect Archive as a Tool for Analyzing Sound Change. Στο P. Austin, H. A. Dry & P. Wittenburg (επιμ.), *Proceedings of the International Workshop on Resources and Tools in Field Linguistics held in conjunction with the Third International Conference on Language Resources and Evaluation*. Las Palmas: Spain, 28.1-28.4.
- Bondi, J. J., Vangsnes, Ø. A., Priestley, J. & K. Hagen. 2014. A multilingual speech corpus of North-Germanic languages. Στο T. Raso & H. Mello (επιμ.), *Spoken corpora and linguistic studies*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 69-83.
- Chambers, J. K. & P. Trudgill. 1980. *Dialectology*. Cambridge: Cambridge University Press.
- Galiotou, E., Karanikolas, N., Manolessou, I., Pantelidis, N., Papazachariou, D., Ralli, A. & G. Xydopoulos. 2014. Asia Minor Greek: towards a computational processing. *Procedia-Social and Behavioral Sciences* 147(Special issue: Proc. IC-ININFO 2013): 458-466.
- Gerner, J., Naramore, E., Owens, M. & M. Warden. 2005. *Professional LAMP: Linux, Apache, MySQL and PHP5 Web Development*. Indianapolis: John Wiley & Sons.
- Glaser, E. 2013. Area formation in morphosyntax. Στο P. Auer, M. Hilpert, A. Stukenbrock & B. Szmrecsanyi (επιμ.), *Space in language and linguistics: geographical, interactional and cognitive perspectives*. Berlin: De Gruyter, 195-221.
- Hymes, D. H. 1962. The ethnography of speaking. Στο T. Gladwin & W. C. Sturtevant (επιμ.), *Anthropology and human behavior*. Washington, D. C.: Anthropology Society of Washington, 13-54.
- Hymes, D. H. 1964. Introduction: Toward ethnographies of communication. *American Anthropologist* 66(6): 1-34.
- Kunst, J. P. & F. Wesseling. 2010. Dialect corpora taken further: The DynaSAND corpus and its application in newer tools. Στο R. Otaguro, K. Ishikawa, H. Umemoto, K. Yoshimoto & Y. Harada (επιμ.), *Proceedings of the 24th Pacific Asia conference on language, information and computation* (Tohoku University, November 4-7, 2010). Waseda University, 759-767.
- Meier, P. et al. 1998. *IDEA: International Dialects of English Archive*. Lawrence: University of Kansas.
- Παπαζαχαρίου, Δ. & Α. Καρασίμος. 2015. Οργάνωση και κωδικοποίηση προφορικών πηγών σε πολυτροπική βάση δεδομένων αιχμής: η περίπτωση του AMiGre corpus. Στο Α. Ράλλη & Σ. Μπόμπολας (επιμ.), *Πρόγραμμα ΘΑΛΗΣ: «Πόντος, Καπαδοκία, Αίβαλι: στα χνάρια της Μικρασιατικής Ελληνικής»*. Πάτρα: Εργαστήριο Νεοελληνικών Διαλέκτων, 55-68.

- Ράλλη, Α. & Σ. Μπόμπολας (επιμ.). 2015. *Πρόγραμμα ΘΑΛΗΣ: «Πόντος, Καππαδοκία, Αϊβαλί: στα χνάρια της Μικρασιατικής Ελληνικής»*. Πάτρα: Εργαστήριο Νεοελληνικών Διαλέκτων.
- Ράλλη, Α., Παπαζαχαρίου, Δ. & Α. Καρασίμος. 2010. Εργαστήριο Νεοελληνικών Διαλέκτων και η βάση δεδομένων Gree.D. Στο Μ. Janse, Β. Joseph, Α. Ralli & Α. Karasimos (επιμ.), *Proceedings of the 4th International Conference on Modern Greek Dialects and Linguistic Theory*. Patras: University of Patras, 7-15.
- Stewart, J. & G. Philipsen. 1984. Communication as situated accomplishment: The cases of hermeneutics and ethnography. *Progress in Communication Sciences* 4: 177-217.
- Wallis, S. & G. Nelson. 2001. Knowledge discovery in grammatically analyzed corpora. *Data Mining and Knowledge Discovery* 5: 305-335.