



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΕΛΟΠΟΝΝΗΣΟΥ

ΥΠΟΛΟΓΙΣΤΙΚΗ ΓΛΩΣΣΟΛΟΓΙΑ

3^η διάλεξη

Π. ΓΑΚΗΣ

ΔΟΜΗ ΠΑΡΟΥΣΙΑΣΗΣ

- Ηλεκτρονικά Λεξικά
- Υπολογιστικά Λεξικά
- Σχεδιασμός **LEXEDIT** (εφαρμογή για υπολογιστικά – μορφολογικά λεξικά)

Ηλεκτρονικά λεξικά

Τα ηλεκτρονικά λεξικά διακρίνονται στα:

- **μηχαναγνώσιμα** (machine readable dictionaries / MRDs), αποτελούν, δηλαδή, μεταφορά εντύπων λεξικών σε αναγνώσιμη μορφή από τον υπολογιστή (σε ταινίες χαρτιού ή μαγνητικές ταινίες). Στα τέλη της δεκαετίας του '60 το **Webster's 7th New Collegiate Dictionary** μεταφέρεται από την έντυπη μορφή με πληκτροδιατηρητήρα (keyrunch) σε ειδική χάρτινη ταινία, αναγνώσιμη από υπολογιστή, με στόχο την υπολογιστική του επεξεργασία, ενώ στη δεκαετία του '70 το **Longman Dictionary of Contemporary English** διατίθεται σε μαγνητική ταινία από τον εκδότη με στόχο την υπολογιστική του επεξεργασία για ερευνητικούς σκοπούς

Ηλεκτρονικά λεξικά

- **λεξικά για ανθρώπους (ΗΛΑ)** τα οποία είτε συντάσσονται με τη βοήθεια επεξεργαστών κειμένου είτε αποτελούν πρωτογενή ΗΛΑ (η πληροφορία τους δηλαδή στηρίζεται σε βάσεις δεδομένων)

ΗΛΕΚΤΡΟΝΙΚΑ ΛΕΞΙΚΑ

- Τα Ηλεκτρονικά Λεξικά παρέχουν όλες τις πολύτιμες για τον εκπαιδευτικό και το μαθητή πληροφορίες (λεξιλογικές, σημασιολογικές και σχετικές με τη χρήση της γλώσσας), τις οποίες παρέχουν και τα παραδοσιακά λεξικά, ενώ ταυτόχρονα τις καθιστούν περισσότερο εύχρηστες και ελκυστικές.
- Τα ηλεκτρονικά λεξικά **συμπληρώνονται με σώματα κειμένων** που παρέχουν στον εκπαιδευτικό ένα ακόμη ευρύτερο και ζωντανό **φάσμα παραδειγμάτων** για τη χρήση των λημμάτων.
- Δυνατή είναι, επίσης, η παράλληλη χρήση των λεξικών

Ηλεκτρονικά Λεξικά

- Τα ηλεκτρονικά λεξικά έχουν σχεδιαστεί έτσι ώστε να μπορεί ο χρήστης είτε
α) να επιλέγει τις πληροφορίες εκείνες που χρειάζεται για κάθε λήμμα είτε
β) να έχει ολόκληρο το άρθρο στη διάθεσή του.

Ηλεκτρονική Λεξικογραφία

- Η ηλεκτρονική λεξικογραφία περιλαμβάνει μια σειρά δραστηριοτήτων που αφορούν τον σχεδιασμό, την κατάρτιση, τη χρήση και την αξιολόγηση ηλεκτρονικών γλωσσικών πόρων με τη χρήση υπολογιστών ή υπολογιστικών εργαλείων. Οι ηλεκτρονικοί γλωσσικοί αυτοί πόροι μπορεί να είναι:
 1. ηλεκτρονικά σώματα κειμένων (ΗΣΚ),
 2. ηλεκτρονικά λεξικά (ΗΛ) ή
 3. υπολογιστικά λεξικά (ΥΛ) ή λεξικά για συστήματα επεξεργασίας φυσικής γλώσσας (ΕΦΓ).

ΗΛΕΚΤΡΟΝΙΚΑ ΛΕΞΙΚΑ

ΛΕΞΙΚΟ
ΤΗΣ ΚΟΙΝΗΣ
ΝΕΟΕΛΛΗΝΙΚΗΣ

ΑΡΧΟΓΡΑΦΗ ΚΑΙ ΠΡΩΤΟΤΥΠΟ ΚΑΙ ΕΚΔΟΣΗ
ΤΟΥ ΚΕΝΤΡΟΥ ΓΙΑ ΝΕΟΕΛΛΗΝΙΚΗΝ ΓΡΑΜΜΑΤΙΚΗΝ
ΕΠΙΤΡΟΠΗΣ ΔΙΑΦΥΛΑΞΕΩΣ ΕΛΛΗΝΙΚΗΣ ΓΛΩΣΣΗΣ


Το Λεξικό της κοινής νεοελληνικής

- Το Λεξικό της κοινής νεοελληνικής είναι ένα σύγχρονο και πλήρες ερμηνευτικό, ορθογραφικό και ετυμολογικό λεξικό της νέας ελληνικής.
- Εκδόθηκε τον Δεκέμβριο του 1998 από το Ινστιτούτο Νεοελληνικών Σπουδών του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης και αποτελεί αποτέλεσμα πολύχρονης και συστηματικής επεξεργασίας.
- Είναι το πρώτο λεξικό της νεοελληνικής που έθεσε επιστημονικές λεξικογραφικές αρχές και επηρέασε με τον τρόπο αυτό καθοριστικά τη σημερινή νεοελληνική λεξικογραφική πραγματικότητα.
- Από τις σημαντικές καινοτομίες του αξίζει να τονιστούν ιδιαίτερα η ύπαρξη φωνητικής μεταγραφής, η σύνδεση κάθε λήμματος με το κλιτικό του παράδειγμα και η προσπάθεια καταγραφής όσο το δυνατόν μεγαλύτερου αριθμού εκφράσεων και φράσεων της νέας ελληνικής. Η εγκυρότητα της ετυμολογικής προσέγγισης είναι επίσης από τα σημαντικότερα πλεονεκτήματά του.


Το Λεξικό της κοινής νεοελληνικής

- **1. Απλή αναζήτηση λήμματος**
- Όταν επιχειρείτε μια **αναζήτηση**, εμφανίζεται ο κυλιόμενος πίνακας (drop down menu) με όλους τους λεξικούς τύπους οι οποίοι ξεκινούν από το γράμμα που έχει πληκτρολογηθεί.
- Όταν επιλέγετε «**Αναζήτηση και στο σώμα των λημμάτων**», στα αποτελέσματα εμφανίζονται όλα τα λήμματα που στο άρθρο τους (ερμηνεία, παραδείγματα χρήσης, ετυμολογία) απαντάται η λέξη ή το τμήμα της λέξης που πληκτρολογείτε.

Παράδειγμα

 **κάλος** ο [kálos] Ο18 : σκλήρυνση του δέρματος κυρίως στα δάχτυλα του ποδιού ή στο πέλμα, που δημιουργείται συνήθ. από την πίεση του παπουτσιού. ΦΡ *πατάω κπ. στον κάλο*, τον θίγω στο πιο ευαίσθητο σημείο του, με λόγια ή με πράξεις. *έχει κάλο στο μυαλό / στον εγκέφαλο*, για κπ. που δεν τον θεωρούμε πολύ λογικό ή ισορροπημένο. || *ρόζος στα δάχτυλα των χεριών: Έβγαλα κάλους από τις δουλειές.* || ~ στις φωνητικές χορδές, σκλήρυνση.

[ιταλ. *callo* -ς]

 **καλοσύνη** η [kalosíni] Ο30α : ANT κακία. **1α.** η ιδιότητα του καλού ανθρώπου, που θέλει το καλό και την ευτυχία του συνανθρώπου του: *Μου φέρθηκε με ~. Μου έδειξε πολλή ~. Η ~ είναι ζωγραφισμένη στο πρόσωπό του.* || *Αν έχεις την ~ / έχεις την ~ να...* τυποποιημένη έκφραση ευγένειας, όταν θέλουμε να ζητήσουμε κάποια εξυπηρέτηση. *~ σας!*, σε κπ. που εκφράζεται επαινετικά για μας ή που μας εξυπηρετεί σε κτ. **β.** ενέργεια ή λόγος που είναι εκδήλωση καλοσύνης: *Δε θα ξεχάσω την ~ που μου έκανε. Κάνει πολλές καλοσύνες, αγαθοεργίες. Κάνε μου την ~ να...* τη χάρη να... **γ.** (πληθ., οικ.) προτερήματα, χαρίσματα: *Αυτό το παιδί έχει πολλές καλοσύνες.* **2.** (παρωχ.) καλός καιρός.

[μσν. *καλοσύνη* < *καλ(ός)* -οσύνη]

Το Λεξικό της κοινής νεοελληνικής

- **2. Αναζήτηση συνόλου λημμάτων**
- Μπορείτε να πληκτρολογήσετε μέρος του λήμματος, χρησιμοποιώντας τον αστερίσκο * για το υπόλοιπο.
- Η επιλογή αυτή είναι χρήσιμη, όταν θέλετε να ζητήσετε ομάδες λέξεων που περιέχουν το μόρφημα ή την ακολουθία χαρακτήρων που πληκτρολογείτε.

Παραδείγματα:

- ✓ κρατ* (όλες οι λέξεις που αρχίζουν από «κρατ»): κράτος, κρατισμός, κρατώ κλπ.
- ✓ *κρατ* (όλες οι λέξεις που περιέχουν το «κρατ»): αισθησιοκράτης, αισθησιοκρατία, αιτιοκρατία κλπ.
- ✓ *ισμός (όλες οι λέξεις που λήγουν σε «ισμός»): ανιμισμός, κρατισμός, υλισμός κλπ.
- ✓ κ*ισμός (όλες οι λέξεις που αρχίζουν από «κ» και λήγουν σε «ισμός»): κλασικισμός, κομμουνισμός, κρατισμός κλπ

Το Λεξικό της κοινής νεοελληνικής

3. Σύνθετη Αναζήτηση

- Στη «**Σύνθετη Αναζήτηση**» μπορείτε να αναζητήσετε λήμματα ή ομάδες λημμάτων από το λεξικό, που εκπληρούν συνδυασμούς κριτηρίων.
- Τα κριτήρια λειτουργούν **συζευκτικά**: **όσα πιο πολλά κριτήρια δώσουμε τόσο λιγότερα λήμματα θα εμφανιστούν**, γιατί ζητάμε λήμματα που να ικανοποιούν ΟΛΑ τα κριτήρια που εισάγονται. Όταν κάποιο κριτήριο δεν μας ενδιαφέρει, το αφήνουμε κενό.

Το Λεξικό της κοινής νεοελληνικής

Παράδειγμα:

• Μπορείτε να βρείτε ΟΥΣΙΑΣΤΙΚΑ που προέρχονται από τα ΙΤΑΛΙΚΑ και έχουν σχέση με τη ΘΑΛΑΣΣΑ.

The screenshot shows the website 'www.greek-language.gr' with the title 'Νέα Ελληνική Γλώσσα'. The main content area is titled 'Λεξικό της κοινής νεοελληνικής' and features a search bar with the text 'Σύνθετη Αναζήτηση' and a 'Βρες' button. Below the search bar, there are several dropdown menus for filtering results. Three red arrows point to the 'Τύπος /λήματος', 'Ερμηνεία', and 'Ετυμολογία' dropdowns. The footer contains the text '© 2006 Κέντρο Ελληνικής Γλώσσας' and 'Διαίκκιο Πνευματικής Σδοκότητας | Όροι Χρήσης'.

Σύνθετη Αναζήτηση

- **Παράδειγμα Χρήσης**

Εδώ πληκτρολογείτε μια λέξη, η οποία περιμένετε να εμφανιστεί σε πρόταση που είναι παράδειγμα χρήσης του λήμματος.

- **Επίπεδο Ύφους**

Το μενού προεπιλογών που εμφανίζεται περιλαμβάνει διαφορετικά επίπεδα ύφους, τα οποία αφορούν είτε όλο το λήμμα είτε κάποια από τις σημασίες του, ώστε να επιλέξετε αυτό που σας ενδιαφέρει. (πχ. Ειρωνικό)

- **Χρήση**

Επιλέγετε την επιθυμητή χρήση της λέξης. Αυτό θα περιορίσει την αναζήτηση σε λέξεις που μία τουλάχιστον σημασία τους έχει τη συγκεκριμένη χρήση.

- **Σημασιολογικό πεδίο**

Επιλέγετε το σημασιολογικό πεδίο στο οποίο ανήκει η σημασία ή κάποια από τις σημασίες του λήμματος. Η επιλογή αυτή είναι χρήσιμη, όταν θέλετε να ζητήσετε ομάδες λέξεων που σχετίζονται με κάποια θεματική περιοχή.

- **Γραμματική - Σύνταξη**

Επιλέγετε τον επιθυμητό τρόπο σύνταξης της λέξης. Αυτό θα περιορίσει την αναζήτηση σε λέξεις που μία τουλάχιστον σημασία τους συντάσσεται με τον συγκεκριμένο τρόπο.

Το Λεξικό της κοινής νεοελληνικής

- https://www.greek-language.gr/greekLang/modern_greek/tools/lexica/triantafyllides/digital_b.html

Αντίστροφο Λεξικό της Ν. Ελληνικής



Αντίστροφο Λεξικό της Ν. Ελληνικής

- Το Αντίστροφο λεξικό είναι πρόσφορο για αναζήτηση λέξεων βάσει της κατάληξής τους. Πληκτρολογώντας μία οποιαδήποτε κατάληξη θα εμφανιστεί η λίστα των λέξεων του λεξικού με την κατάληξη αυτή. Ισχύουν οι ίδιοι κανόνες πληκτρολόγησης λημμάτων με τη διαφορά ότι δεν απαιτείται η χρήση του * στην αρχή. Για παράδειγμα, πληκτρολογώντας "ισμός" θα εμφανιστούν οι ίδιες λέξεις με το εάν πληκτρολογήσετε "*ισμός". Μπορείτε ωστόσο να χρησιμοποιήσετε το * για πιο περίπλοκα πρότυπα καταλήξεων.

Αντίστροφο Λεξικό της Ν. Ελληνικής

- **1. Αναζήτηση λήμματος**

Το πρώτο χαρακτηριστικό του λεξικού που συναντάτε όταν επιχειρείτε μια αναζήτηση είναι ο "**κυλιόμενος πίνακας**" (**drop-down**) με όλους τους λεξικούς τύπους που ξεκινούν από το γράμμα που έχει πληκτρολογηθεί.

Επιλέγοντας οποιοδήποτε γράμμα του ελληνικού αλφαβήτου μεταφερόσατε στην οθόνη όπου παρατίθενται αλφαβητικά όλα τα λήμματα του αντίστοιχου γράμματος του λεξικού.

Εμφανίζεται επίσης η πρόσθετη επιλογή «Αποστολή στα Σώματα» κάτω από τον τίτλο «Αναζήτηση». Μέσω αυτής, έχετε τη δυνατότητα να επιλέξετε για ένα συγκεκριμένο λήμμα, προϊόν κάποιας μορφής αναζήτησης στα λεξικά, να δείτε τις εμφανίσεις του στα σώματα κειμένων.

Όταν επιλέγετε "Αναζήτηση και στο σώμα των λημμάτων", στα αποτελέσματα εμφανίζονται όλα τα λήμματα όπου απαντάται η λέξη ή το τμήμα λέξης που πληκτρολογείτε

Αντίστροφο Λεξικό της Ν. Ελληνικής

- **2. Αναζήτηση συνόλου λημμάτων**

Ένα ακόμη χρήσιμο χαρακτηριστικό που υπάρχει ενσωματωμένο στο λεξικό είναι η χρήση του συμβόλου "*". Επιτρέπει μια "διαφορετική" αναζήτηση ενός λήμματος. Η επιλογή αυτή είναι χρήσιμη, όταν θέλετε να ζητήσετε ομάδες λέξεων που περιέχουν το μόρφημα ή την ακολουθία χαρακτήρων που πληκτρολογείτε.

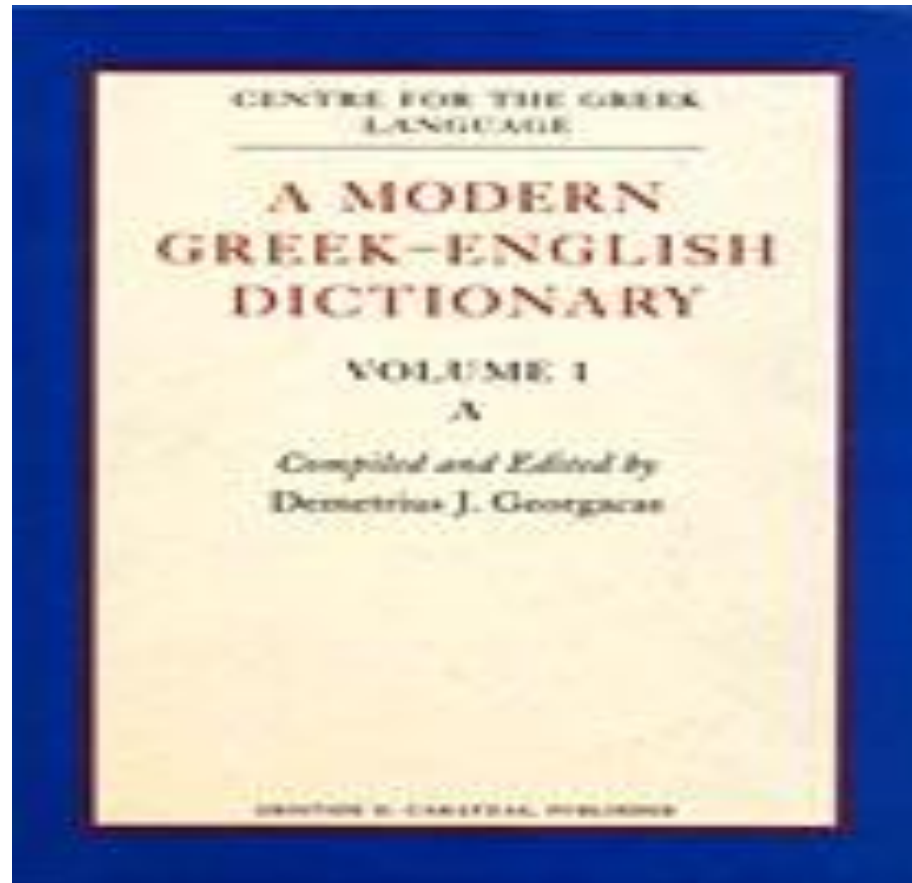
*Αν για παράδειγμα χρησιμοποιήσετε "*παρα*", θα βρείτε λήμματα που περιέχουν στο ερμήνευμά τους το "παρα" σε οποιαδήποτε θέση, είτε στην αρχή, στη μέση ή στο τέλος αυτού, όπως τα "παραξενάδα", "απαράλλαγα", "αμπάρα".*



Αντίστροφο Λεξικό της Ν. Ελληνικής

- https://www.greek-language.gr/greekLang/modern_greek/tools/lexica/reverse/index.html

Ελληνο-αγγλικό Λεξικό (Γεωργακά)



Ελληνο-αγγλικό Λεξικό (Γεωργακά)

- Το Modern Greek-English Dictionary του Δ. Γεωργακά, είναι, στην πραγματικότητα, το εθνικό λεξικό της Νέας Ελληνικής γλώσσας, όπου: **για κάθε λέξη διακρίνονται σχολαστικά οι σημασίες και τα επίπεδα χρήσης και δίνεται η ιστορία και η ετυμολογία της.** Προπάντων όμως παρέχεται –για πρώτη φορά στα χρονικά της νεοελληνικής λεξικογραφίας– πλήρης **τεκμηρίωση** για όλα αυτά, μέσα από χωρία, αποδελτιωμένα από αυθεντικές γραμματειακές πηγές της Νέας Ελληνικής (από το 1800 έως σήμερα).
- Το λεξικό περιλαμβάνει, προς το παρόν, μόνο το στοιχείο Α (α-άωτον)

 **αγαθός, -ή, -ό** [αγαθός] ① **honest and virtuous, fine, benevolent, good** (syn καλός, καλοκάγαθος, ant κακός, πονηρός): ~ άνθρωπος good man | αγαθή ψυχή a good soul | αγαθό πλάσμα | αγαθή γριούλα | αγαθά μάτια | κάποιος ~ δαίμονας... που με συμπαθεί (Palam) | ένας αγαθότατος ποιητής, ο Πολέμης (id.) | καθίζουν στο скаμνί τους αγαθούς και τους αθώους (id.) they indict people that are virtuous and innocent | το πνεύμα μιας θεότητας αγαθής (Panagiotop) | roem τα γερατειά στην ήσυχη την κοίτη | δεν σε τρομάζανε, αγαθέ πρεσβύτη (Zengoli) ② **good, well-meaning**: έχουμε αγαθές σχέσεις we are on good terms | αγαθή βούληση | αγαθή διάθεση | αγαθή προαίρεση | αγαθές πρωτοβουλίες | έχει αγαθές προθέσεις | αγαθή πράξη | αγαθή έκβαση | αγαθή επίδραση | αγαθή εντύπωση | αγαθές συνέπειες | αγαθές εξαιρέσεις | ~ κύκλος versus φαύλος κύκλος (Parapoutsos) | ασκούμε αγαθή διοίκηση απέναντι στο λαό (Panagiotop) | είναι αγαθοί οι σκοποί σου ή κάτι σκοτεινό μηχανεύεσαι; (Theotokas) | γέρος... με αγαθό ύφος, ευπαρουσίαστος (Melas) | να έχη μια καλή και αγαθή ζωή (Kanellor) ③ **agreeable, kind**: αγαθό χαμόγελο a kindly smile | σπανιότατα από τέτοιες σταδιοδρομίες απομένει αγαθή μνήμη (Panagiotop) | roem να, του λύχνου είν' αγαθότερο το φέγγος (Palam) | αλλά τα ξόρκια τ' αγαθά, τις ρητορείες,... | τι θα τα κάνης; (Seferis) ④ **advantageous, beneficial, lucky**: κάτω από την ευλογία ενός αγαθού ανοιξιάτικου ήλιου (Nirvanas) | έχω την αγαθή τύχη να + subj | have the good luck to | phr τύχη αγαθή (L) w. good luck, w. good fortune (syn με το καλό) ⑤ **too kind, naive, simple-minded, credulous** (syn απλοϊκός, αφελής, ανόητος, κουτός): είναι ~ he is naive | αγαθό παιδί simpleton [fr MG ← AG αγαθός]

Ελληνο-αγγλικό Λεξικό (Γεωργακά)

- https://www.greek-language.gr/greekLang/modern_greek/tools/lexica/georgakas/index.html

Ηλεκτρονικά Λεξικά

- Κλιτικό Λεξικό Νέας Ελληνικής (Lexigram)
- Κλιτικό Λεξικό Αρχαίας Ελληνικής (Lexigram)
- Λεξικό της Αρχαίας Ελληνικής Liddell-Scott
- Κλιτικό Λεξικό Αρχαίας Ελληνικής (Lexigram)
- Αντίστροφο Λεξικό (Αναστασιάδη – Συμεωνίδη)
- Λεξικό Συνωνύμων και Αντωνύμων (Lexigram)
- Λεξικό Γλωσσολογικών Όρων
- Λεξικό Επιλεγμένων Λέξεων και Φράσεων
- Λεξικό Ομορρίζων και Ετυμολογίας (Lexigram)
- Ερμηνευτικό και Ελληνοαγγλικό Λεξικό Δημοτικού (Lexigram)
- Λεξικό Σουίδα
- Λεξικό Λατινικών (Lexigram)
- Ελληνοαγγλικό λεξικό Liddell-Scott
- Ελληνοαγγλικό Λεξικό (Δ. Γεωργακά)
- Αγγλοελληνικό/Ελληνοαγγλικό Λεξικό
- Ελληνορωσικό Λεξικό

Υπολογιστικό Λεξικό

- φέρει πληροφορία: α) ορθογραφική (ορθή γραφή του κλιτικού τύπου),
- β) μορφηματική (το είδος των μορφημάτων: πρόθημα, θέμα, επίθημα, κατάληξη, που απαρτίζουν τον κλιτικό τύπο),
- γ) μορφοσυντακτική (μέρος του λόγου, γένος, πτώση, πρόσωπο κ.λπ.),
- δ) υφολογική (τα υφολογικά χαρακτηριστικά του τύπου: προφορικό, λόγιο κ.λπ.) και
- ε) ορολογική (επιπλέον πληροφορία για το αν ο τύπος αποτελεί μέρος ειδικού λεξιλογίου)

Μορφολογικά υπολογιστικά λεξικά της Ν. Ελληνικής

- Ένας από τους πιο διαδεδομένους αλγόριθμους μορφολογικής επεξεργασίας είναι το **PC-KIMMO** (Koskenniemi, 1983), το μορφολογικό μοντέλο των δύο επιπέδων του Koskenniemi που ονομάστηκε έτσι, από τον εμπνευστή του, Kimmo Koskenniemi.
- Το PC-KIMMO αποτελεί κατ' ουσία ένα πρόγραμμα σχεδιασμένο να δημιουργεί (generate), αλλά και να αναγνωρίζει (recognize) λέξεις, και χρησιμοποιεί ένα μοντέλο δύο επιπέδων για τη δομή της λέξης.
- Το μοντέλο δύο επιπέδων του Koskenniemi αφορά αφενός **στους τρόπους διάταξης των μορφημάτων** (morphotactics) και αφετέρου στις **μορφοφωνολογικές μεταβολές** (morphophonemics) (Detorakis, 2009).

Μορφολογικά υπολογιστικά λεξικά της Ν. Ελληνικής

- Οι (Sgarbas, Fakotakis & Kokkinakis, 1995) έχουν δημιουργήσει ένα αντίστοιχο μοντέλο δύο επιπέδων για τη μορφολογική περιγραφή της νέας ελληνικής.
- Το υλοποιημένο σύστημα καλύπτει τις ακόλουθες κατηγορίες: ουσιαστικά, επίθετα, ομαλά ρήματα καθώς και μετοχές ως εγκλίσεις ρημάτων.
- Χρησιμοποιεί **36 κανόνες** για τον χειρισμό των διφθόγγων, κανόνες για τη μεταβολή του τονισμού, τον συλλαβισμό, τη ρηματική αύξηση και άλλες ιδιαιτερότητες της ελληνικής μορφολογίας.

Υπολογιστικά Λεξικά

- Για την ελληνική γλώσσα διατίθεται επιπλέον:
 1. το υπολογιστικό μορφολογικό και συντακτικό λεξικό της Ν. Ελληνικής που αναπτύχθηκε από το **ΙΕΛ**. Περιλαμβάνει 20.149 λήμματα κωδικοποιημένα σε μορφολογικό και συντακτικό επίπεδο σύμφωνα με το μοντέλο PAROLE.
 2. Το ανεξάρτητο εργαλείο το **Υπολογιστικό Μορφολογικό Λεξικό της νέας ελληνικής γλώσσας** (Baltzis, Kolalas & Eumeridou, 2005b).

Υπολογιστικό μορφολογικό και συντακτικό λεξικό της Νέας Ελληνικής (ΙΕΛ)

- Περιλαμβάνει **20.149** λήμματα κωδικοποιημένα σε **μορφολογικό** και **συντακτικό** επίπεδο, σύμφωνα με το μοντέλο PAROLE, που βασίζεται στα διεθνή γλωσσολογικά πρότυπα.
- Με το ίδιο μοντέλο έχουν αναπτυχθεί λεξικά για 11 ακόμη ευρωπαϊκές γλώσσες (Αγγλικά, Γαλλικά, Γερμανικά, Δανικά, Ισπανικά, Ιταλικά, Καταλανικά, Ολλανδικά, Πορτογαλικά, Σουηδικά, Φινλανδικά).

Περιεχόμενα λεξικού ΙΕΛ

- Για την επιλογή των 20.000 λημμάτων που περιλαμβάνονται στο λεξικό ακολουθήθηκε μια υβριδική διαδικασία:
 1. στατιστική επεξεργασία ενός Σώματος Κειμένων 9.000.000 λέξεων περίπου, με στόχο τον εντοπισμό των συχνότερων λημμάτων,
 2. επεξεργασία του καταλόγου των συχνότερων λημμάτων με βάση γλωσσολογικά κριτήρια.
- Πιο συγκεκριμένα, το λεξικό περιλαμβάνει:
 1. 20.149 μορφολογικές μονάδες, και
 2. 25.092 συντακτικές μονάδες (κωδικοποιείται, με τη μορφή συντακτικών μονάδων, η **συντακτική συμπεριφορά του λήμματος**: κωδικοποιούνται, δηλαδή, τα συμπληρώματα με τα οποία συντάσσεται ένα λήμμα, καθώς και τα χαρακτηριστικά που απαιτούνται για τον χαρακτηρισμό και την αναγνώριση των συμπληρωμάτων αυτών (π.χ. αν πρόκειται για υποκείμενο - ουσιαστικό σε πτώση ονομαστική, κτλ.).).

Ειδικά Υπολογιστικά Λεξικά

- Στα ειδικά ηλεκτρονικά λεξικά για την ελληνική γλώσσα μπορούν να αναφερθεί το

ηλεκτρονικό λεξικό

οικονομικής - χρηματιστηριακής ορολογίας.

- Για τη δημιουργία του λεξικού αυτού χρησιμοποιήθηκαν τα προγράμματα κωδικοποίησης και αυτόματης κλίσης τα οποία έχουν χρησιμοποιηθεί για την κατασκευή των γενικών ηλεκτρονικών λεξικών της νέας ελληνικής από τη Μονάδα Αυτόματης Επεξεργασίας Φυσικών Γλωσσών (<http://linginfo.frl.auth.gr>) του Εργαστηρίου Μετάφρασης και Επεξεργασίας του Λόγου (Ε.Μ.Ε.Λ.) του Α.Π.Θ. (Tziafa, 2007).
- **Αδυναμία**: περιορισμένος όγκος τους και ο εξειδικευμένος χαρακτήρας τους που περιγράφει συγκεκριμένα επίπεδα της γλώσσας

Υπολογιστική Λεξικογραφία

Κέρδη:

- Γνώση σχετικά με μεμονωμένες λέξεις μιας γλώσσας: απαραίτητη για κάθε είδους επεξεργασία φυσικής γλώσσας
- Συστήματα Μηχανικής Μετάφρασης: από τις πρώτες εφαρμογές υπολογιστικής λεξικογραφίας Εφαρμογές Speech-to-Text και Text-to-Speech
- Ηλεκτρονικά λεξικά τσέπης: Εργαλεία ελέγχου ορθογραφίας
- **Κοινό αίτημα**: σχεδιασμός λεξικών πηγών ενιαίας μορφής έτσι ώστε να διευκολύνεται η ευρεία χρήση τους και να αποφεύγεται η επανάληψη των διαδικασιών ανάπτυξής τους

Χρήση Η/Υ στην Λεξικογραφία

Πώς;

1. Συλλογή γλωσσικών δεδομένων
2. Αυτοματοποίηση διαδικασιών γλωσσολογικής έρευνας
3. Υπολογιστικές εφαρμογές αποτελεσμάτων γλωσσολογικής έρευνας (Garvin 1962)

Πού;

1. Υπολογιστική παραγωγή βοηθημάτων για γλωσσολογική (και λογοτεχνική) ανάλυση
2. Γλωσσική αυτοματοποίηση - ανάλυση & σύνθεση
3. Προσομοίωση της γλωσσικής δυναμικής (Καινοτόμες τεχνολογίες και εικονικά περιβάλλοντα διδασκαλίας στη γλωσσική εκπαίδευση)
4. Ανάκτηση πληροφοριών
5. Στατιστική ανάλυση γλωσσικών φαινομένων (Lamb1961)

Υπολογιστικά Λεξικά

- Τα υπολογιστικά λεξικά είτε:
 1. αποτελούν απαραίτητο στοιχείο για κάθε σύστημα επεξεργασίας φυσικής γλώσσας (ΕΦΓ) είτε
 2. είναι αναγνώσιμα από ανθρώπους, όπως οι λεξιλογικές βάσεις δεδομένων του Wordnet και του Framenet

Χαρακτηριστικά Υπολογιστικών Λεξικών

- Χαρακτηριστικό στοιχείο των υπολογιστικών λεξικών είναι ο **απεριόριστος χώρος αποθήκευσης δεδομένων** όπως επίσης και η **δυνατότητα αποθήκευσης περισσότερων πληροφοριών σε περισσότερα επίπεδα**. *The absence of space constraints call for more, not less intellectual discipline in the selection and arrangement of information'*(Hanks, 2001).

Χαρακτηριστικά Υπολογιστικών λεξικών

- Υπάρχει επιπλέον **μεγαλύτερη ταχύτητα** κατά την ανάκληση, επεξεργασία και παρουσίαση των δεδομένων, **δυνατότητα ταυτόχρονης αναζήτησης σε πολλά λεξικά ταυτόχρονα** όπως επίσης και **δυνατότητα συχνότερης συντήρησης και επικαιροποίησης** σε σχέση με τα έντυπα.

Χαρακτηριστικά Υπολογιστικών Λεξικών

- Ένα ακόμη χαρακτηριστικό είναι η διαφοροποίηση της προσβασιμότητας στη μακροδομή (πρόταση) σε σχέση με τα έντυπα λεξικά.
- Η απλή αναζήτηση στον αλφαβητικό κατάλογο των εισόδων λημμάτων κατά τη λογική του έντυπου λεξικού (ταίριασμα συμβολοσειράς) διατηρείται και στα ΗΛΑ.
- Υπάρχει όμως και η δυνατότητα ενσωμάτωσης αλγορίθμου που ενσωματώνεται στους ορθογραφικούς διορθωτές (fuzzy matching). Τμήματα των κειμένων αυτών ανασύρονται με απόλυτη (full match) ή μερική (fuzzy match) ταύτιση



Ευχαριστώ για την προσοχή σας!