



# Τη γλώσσα μου έδωσαν ελληνική

Μαρία Γαβριηλίδου, ΙΕΛ/ΕΚ Αθηνά

Παγκόσμια Ημέρα Μετάφρασης

29 Σεπτεμβρίου 2018, ΕΙΕ

# Πρώτη τοποθέτηση

Η ΓΛΩΣΣΑ ΤΩΝ ΝΕΩΝ

# Η γλώσσα των νέων

- ▶ όχι ενιαίο και αυτοτελές γλωσσικό σύστημα
- ▶ συνδέει το δίκτυο των συνομηλίκων → κοινωνιόλεκτος
- ▶ χρησιμοποιείται σε ορισμένες συνθήκες επικοινωνίας
- ▶ η χρήση της ορίζει μια κοινότητα
- ▶ σύνδεση με γεωγραφικές και κοινωνικές παραμέτρους
- ▶ έχει λεξιλογικά, συντακτικά και πραγματολογικά χαρακτηριστικά
- ▶ πάντα υπήρχε και πάντα θα υπάρχει!

# Χαρακτηριστικά της γλώσσας των νέων

- ▶ εξαιρετικά δημιουργική και παραγωγική
- ▶ τροφοδοτεί πρότυπη γλώσσα με νέες λέξεις
- ▶ όσο παραμένει «κλειστή» έχει συνθηματικό χαρακτήρα (αναγνωριστικό κοινότητας)
- ▶ διαδικασίες δημιουργίας και ανανέωσης λεξιλογίου
  - ▶ νέες σημασίες σε υπάρχουσες λέξεις [μπάζο, κόκαλο]
  - ▶ νέες συντακτικές δομές [ψήνεις να πάμε συναυλία; τι φάση;]
  - ▶ δανεισμός από άλλες γλώσσες με ή χωρίς εξελληνισμό μέσω κατάληξης [ποζεράς / ποζερού / ποζέρι]
  - ▶ αναβίωση παλαιότερων λέξεων / εκφράσεων με ίδια ή και νέες σημασίες [νισάφι πια!]
- ▶ ραγδαίες αλλαγές / ανανέωση εκφραστικών μέσων (όταν κάτι ενσωματωθεί στην πρότυπη γλώσσα εγκαταλείπεται γιατί παύει να εξυπηρετεί τον συνθηματικό ρόλο της κοινωνιολέκτου)
- ▶ άρρηκτα συνδεδεμένη με την ηλικία → σε ατομικό επίπεδο παύει να χρησιμοποιείται όταν ο χρήστης «ενηλικιωθεί» (ένταξη σε εργασιακό περιβάλλον, επίσημο επίπεδο ύφους)



# Δεύτερη τοποθέτηση

Η ΕΛΛΗΝΙΚΗ ΓΛΩΣΣΑ ΣΤΗΝ ΨΗΦΙΑΚΗ ΕΠΟΧΗ

# Η ελληνική γλώσσα στην ψηφιακή εποχή

- ▶ αν μία γλώσσα δεν υπάρχει στα ψηφιακά μέσα κινδυνεύει
- ▶ γιατί κινδυνεύει;
  - ▶ προφορική γλώσσα > γραπτή γλώσσα > ψηφιακή γλώσσα
    - ▶ *The digital revolution is comparable to Gutenberg's invention of the printing press.*
    - ▶ *While the printing press helped step up the exchange of information in Europe, it also led to the extinction of many languages. Regional and minority languages were rarely printed and languages such as Cornish and Dalmatian were limited to oral forms of transmission, which in turn restricted their scope of use. Will the Internet have the same impact on our modern languages?*

# αναγκαία μέσα για την ψηφιακή επιβίωση μιας γλώσσας

- ▶ ψηφιακά κείμενα / πόροι
- ▶ εργαλεία επεξεργασίας τους
- ▶ ψηφιακά εγγράμματοι χρήστες
- ▶ πολιτική ενίσχυσης και έμπρακτης υποστήριξης του ψηφιακού γραμματισμού και της έρευνας στη γλωσσική τεχνολογία



## filter by:

► Language

▼ Resource Type

- ◆ Corpus (404)
- ◆ Lexical Conceptual Resource (46)
- ◆ Tool Service (35)

► Media Type

► Availability

► Licence

► Restrictions of Use

► Validated

► Foreseen Use

► Use Is NLP Specific

► Linguality Type

► Multilinguality Type

► Modality Type








► MIME Type

► Conformance to Standards/Best

## 485 language resources (page 1 of 25)

Number of Downloads ▾

Next »

- ACCURAT balanced test corpus for under resourced languages**  ↓ 0 👁 44  
Croatian English Estonian German Greek, Modern (1453-) Latvian Lithuanian Romanian Slovenian
- ACCURAT corpus of comparable sentences**  ↓ 0 👁 32  
Croatian English Estonian German Greek, Modern (1453-) Latvian Lithuanian Romanian Slovenian
- aformes**  ↓ 0 👁 44  
English Greek, Modern (1453-)
- Alternate Minister for Infrastructure, Transport, and Networks Michalis Papadopoulos' various documents**  ↓ 0 👁 21  
Greek, Modern (1453-)
- Antonis Manidakis' various documents**  ↓ 0 👁 29  
Greek, Modern (1453-)
- A parallel corpus collected from the European Constitution**  ↓ 0 👁 33  
Czech Danish Dutch; Flemish English Estonian Finnish French German Greek, Modern (1453-) Hungarian Irish Italian Latvian Lithuanian Maltese Polish Portuguese Slovak Slovenian Spanish; Castilian Swedish
- A parallel corpus of KDE4 localization files (v.2)**  ↓ 0 👁 35  
Afrikaans Arabic Armenian Assamese Asturian; Bable; Leonese; Asturleonese Basque Belarusian Bengali Bokmål, Norwegian; Norwegian Bokmål Breton Bulgarian Catalan; Valencian Central Khmer Chinese Crimean Tatar; Crimean Turkish Croatian Czech Danish Dutch; Flemish English Esperanto Estonian Finnish French Galician Georgian German Greek, Modern (1453-) Gujarati Hausa Hebrew Hindi Hungarian Icelandic Indonesian Irish Italian Japanese Korean Latvian Lithuanian Maltese Maldivian Malay Maltese Marathi Meitei Mayek Mizo Myanmar Nepali Norwegian Persian Polish Portuguese Romanian Russian Serbian Slovenian Spanish; Castilian Swedish



# εργαλεία

- ▶ μετατροπείς από μορφότυπο σε μορφότυπο
  - ▶ ορθογραφικοί, συντακτικοί κ.ά. διορθωτές
  - ▶ δομικοί αναλυτές (παράγραφο, πρόταση, λέξη, ...)
  - ▶ γλωσσικοί αναλυτές (λημματοποίηση, μορφολογική, συντακτική, σημασιολογική ανάλυση, αναγνώριση γνώμης/συναισθήματος ομιλητή...)
  - ▶ εργαλεία εξόρυξης πληροφορίας/γνώσης
  - ▶ εργαλεία αναγνώρισης ονοματικών οντοτήτων
- 
- ▶ μονόγλωσσα / δίγλωσσα / πολύγλωσσα → μεταφραστικά εργαλεία

# Ποιότητα μεταφραστικής τεχνολογίας για 30 γλώσσες

Excellent support	Good support	Moderate support	Fragmentary support	Weak/no support
	English	French Spanish	Catalan Dutch German Hungarian Italian Polish Romanian	Basque Bulgarian Croatian Czech Danish Estonian Finnish Galician <b>Greek</b> Icelandic Irish Latvian Lithuanian Maltese Norwegian Portuguese Serbian Slovak Slovene Swedish

# Υπαρξη τεχνολογιών επεξεργασίας κειμένου σε 30 γλώσσες

Excellent  
support



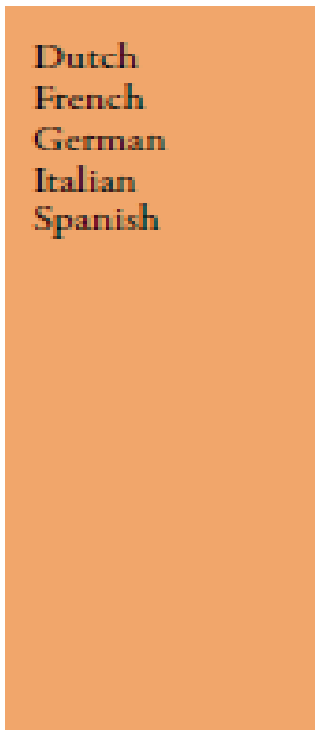
Good  
support

English



Moderate  
support

Dutch  
French  
German  
Italian  
Spanish



Fragmentary  
support

Basque  
Bulgarian  
Catalan  
Czech  
Danish  
Finnish  
Galician  
Greek  
Hungarian  
Norwegian  
Polish  
Portuguese  
Romanian  
Slovak  
Slovene  
Swedish

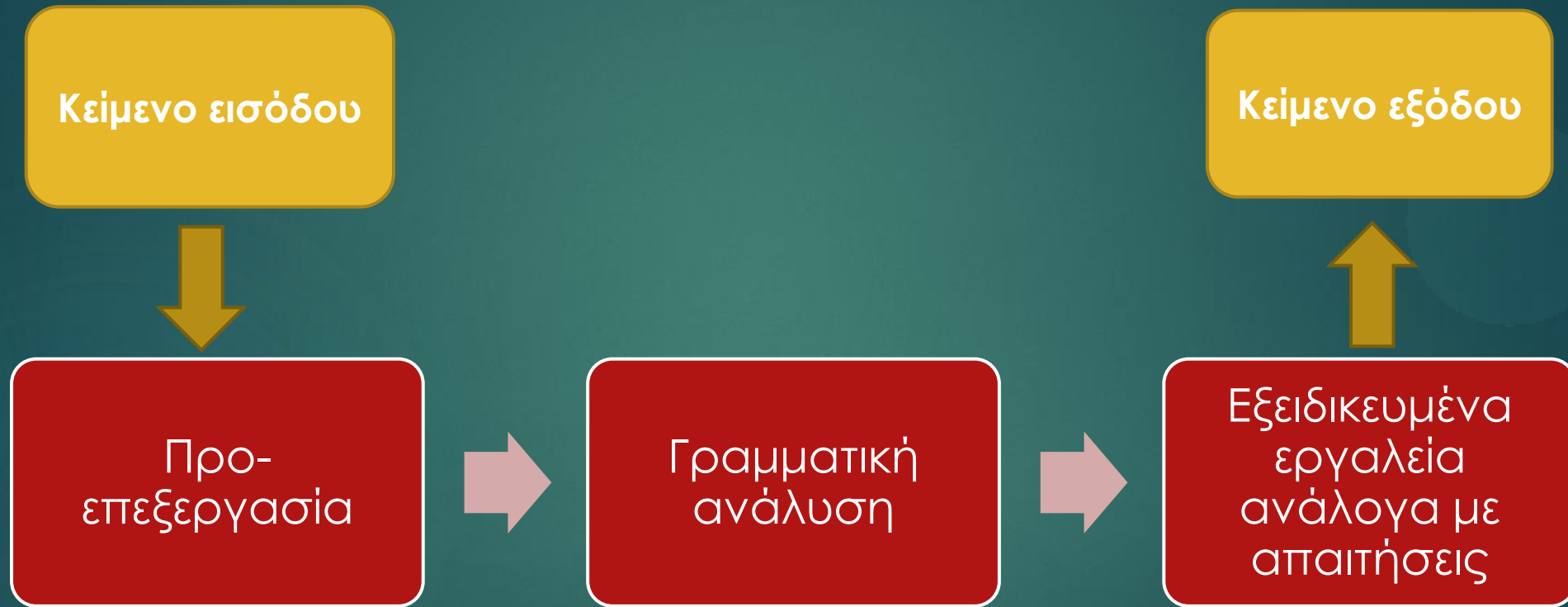


Weak/no  
support



Croatian  
Estonian  
Icelandic  
Irish  
Latvian  
Lithuanian  
Maltese  
Serbian



# Βασική αλυσίδα επεξεργασίας κειμένου



# Ψηφιακά εργαλεία επεξεργασίας ΝΕ







clarin:el  



## browse resources










filter by:

- Resource Type
  - Tool Service (35)
    - Tool/Service Type
    - Tool/Service Subtype
    - Language Dependent
    - InputInfo/OutputInfo Resource Type
    - InputInfo/OutputInfo Media Type
    - Annotation Type
    - Annotation Format
    - Evaluated
  - Language
  - Media Type
  - Availability
  - Licence

### 35 language resources (page 1 of 2)

-  **GrNE-Tagger**  
Greek, Modern (1453-)
-  **ILSP Lemmatizer**  
Greek, Modern (1453-)
-  **Text Summarization Tool**  
Greek, Modern (1453-)
-  **Sentiment Analysis Tool**  
Greek, Modern (1453-)
-  **ILSP Language Identification System**  
Dutch; Flemish English French German Greek, Modern (1453-)
-  **Word frequency counter**  
Greek, Modern (1453-)

- Greek, Modern (1453-)
-  **ILSP Sentence splitter and tokenizer for Greek text**  
Greek, Modern (1453-)
-  **ILSP Term Extractor**  
Greek, Modern (1453-)
-  **ILSP Dependency parser**  
Greek, Modern (1453-)
-  **ILSP Feature-based multi-tiered POS Tagger**  
Greek, Modern (1453-)
-  **HTokenizer**  
Greek, Modern (1453-)
-  **Greek Texts Feature Extractor**
-  **COMBO - Tool for extracting words based on a spelling or pronunciation criteria**  
Greek, Modern (1453-)
-  **HGazetteer**  
Greek, Modern (1453-)
-  **OpenNLP Sentence Detector (Portuguese)**  
Portuguese

# Σύστημα μεταγραφής μεταξύ Ελληνικών και Greeklish



All Greek to me!<sup>TM</sup>  
Any Greeklish to Greek  
Online Demo Page



[Αρχική Σελίδα](#)  
[IEA](#)

[Αρχική Σελίδα](#)  
[All Greek to me!](#)

[Σελίδα επίδειξης](#)  
[All Greek to me!](#)

[Νέα](#)

[Υποστήριξη](#)

[Λύσεις](#)

[Επικοινωνία](#)

[Συντελεστές](#)

Το σύστημα επίδειξης σε αυτήν την σελίδα επιτρέπει στον χρήστη να χρησιμοποιήσει το αυτόματο σύστημα μετατροπής από Greeklish σε ελληνικά, για μη εμπορικούς σκοπούς. Η χρήση του περιορίζεται στους 2000 χαρακτήρες.

Greeklish		Greek
th glwssa moy edwsan ellhnikh	→	τη γλώσσα μου έδωσαν ελληνική
ti glossa mou edosan elliniki		τι γλώσσα μου έδωσαν ελληνική
thelo na grapso ellinika		θέλω να γράψω ελληνικά
8elo na grapsw ellhnika		θέλω να γράψω ελληνικά
xerw na grafw agglika	←	ξέρω να γράφω αγγλικά
ksero na grafo anglika		ξέρω να γράφω αγγλικά

Μέχρι τώρα έχουν μετατραπεί 125,449,305 λέξεις!

Θέλετε να μας πείτε την γνώμη σας; [Επικοινωνήστε μαζί μας!](#)

# Σύστημα μεταγραφής μεταξύ Ελληνικών και Greeklish – ΕΛ → GL



All Greek to me!<sup>TM</sup>  
Any Greeklish to Greek  
Online Demo Page



[Αρχική Σελίδα  
IEL](#)

[Αρχική Σελίδα  
All Greek to me!](#)

[Σελίδα επίδειξης  
All Greek to me!](#)

[Νέα](#)

[Υποστήριξη](#)

[Λύσεις](#)

[Επικοινωνία](#)

[Συντελεστές](#)

Το σύστημα επίδειξης σε αυτήν την σελίδα επιτρέπει στον χρήστη να χρησιμοποιήσει το αυτόματο σύστημα μετατροπής από Greeklish σε ελληνικά, για μη εμπορικούς σκοπούς. Η χρήση του περιορίζεται στους 2000 χαρακτήρες.

Greeklish		Greek
ti glwssa mou edwsan elliniki	→	τη γλώσσα μου έδωσαν ελληνική
thelw na grapsw ellinika		θέλω να γράψω ελληνικά
kserw na grafw agglika	←	ξέρω να γράφω αγγλικά

Μέχρι τώρα έχουν μετατραπεί 125,567,886 λέξεις!

Θέλετε να μας πείτε την γνώμη σας; [Επικοινωνήστε μαζί μας!](#)