

## An Early Modern Greek Corpus

Eleni Karantzola  
University of the Aegean  
karantzola@rhodes.aegean.gr

Konstantinos Sampanis  
University of the Aegean  
konstantinos.sampanis@yahoo.com

The *Thesaurus Linguae Graecae* (University of California, Irvine) has collected and digitized most literary texts written in Greek from Homer to the fall of Constantinople in 1453 AD. Post-Byzantine or Early Modern Greek (EMG) texts are not systematically included in this corpus, although it is generally accepted that the 16<sup>th</sup> century constitutes an important turn-point concerning the ecology of communication in Europe, as far as the dynamic appearance of vernacular languages — among which EMG- as languages of “literacy” is concerned. Regarding the Greek language, after a poorly documented period (6<sup>th</sup> — 12<sup>th</sup> c.), there is nothing quantitatively and qualitatively comparable to the production of prose texts in vernacular Greek, attested from the late 15<sup>th</sup> century. Classicizing Greek does not anymore dominate in writing, but on the other hand EMG is growing rapidly as a literary language, with aesthetic function, as well as a simple *scripta*, namely a written vernacular for legal, administrative, commercial and other functions.

During the last decade a large number of prose texts in — more or less — vernacular Greek have been published. Our research group at the University of the Aegean (Rhodes), which thoroughly investigates the Early Modern Greek language, have digitized a number of them, in order to create a representative database for the study of the distribution of antagonistic forms and the reconstruction of this important period in the history of the Greek language. Our corpus includes: (i) a selection of over two hundred extracts from vernacular prose of the 16<sup>th</sup> century (cf. Kakoulidou-Panou, Karantzola & Tiktopoulou, in press), which belong to different genres, e.g. chronicles and biographies, lives of saints, homiletic texts, didactic moralizing stories, prophetic and apocalyptic visions, fictions, dictionaries, etc., (ii) two extensive narrative texts, the first composed by a Corfiot in the first half of 16<sup>th</sup> c. and printed in Venice (cf. Kakoulidi-Panou, ed. 2000), and the second one written by a Cretan priest at the end of the century (cf. Kakoulidi-Panou, E. Karantzola & M. Chalvatzidaki, eds. 2007), (iii) a number of notary books, especially from the Ionian islands, Crete, and other regions of Greece (archives mainly in scanned form from the original).

In our presentation we will discuss the main characteristics of the corpus, methodological issues concerning the data and the new perspectives of corpus-based research for Historical Linguistics and the research of the diachrony of the Greek language.

### References

- Kakoulidi-Panou, E. ed. 2000. *Ioannikios Kartanos, Palaia te kai Nea Diathiki* (Venice 1536) [in Greek]. Thessaloniki: Centre for Greek Language.
- Kakoulidi-Panou, E., E. Karantzola & M. Chalvatzidaki, eds. 2007. *Ioannis Morezinos, Klini Solomontos* (Venice: 1599) [in Greek]. Irakleio: Etairia Kritikou Istorikon Meleton.
- Kakoulidi-Panou, E., E. Karantzola & K. Tiktopoulou (in press). *Demotic prose texts of the 16<sup>th</sup> c.* [in Greek]. Thessaloniki: Centre for Greek Language.