

ICININFO

Asia Minor Greek: Towards a Computational Processing

Eleni Galiotou^{a*}, Nikitas Karanikolas^a, Ioanna Manolesou^b, Nikolaos Pantelidis^b,
Dimitris Papazachariou^c, Angela Ralli^c, George Xydopoulos^c

^a*Dept. of Informatics, Technological Educational Institute of Athens, Ag. Spyridonos, Egaleo Athens GR-122 10, Greece*

^b*Dept. of Philology, University of Athens, Panepistimiopolis, Ilisia Athens GR-116 35, Greece*

^c*Dept. Of Philology, University of Patras, Panepistimiopolis, Rio Patras GR-265 04, Greece*

Abstract

In this paper, we discuss issues concerning the computational aspect of an on-going research project which aims at providing a systematic study of three Greek dialects of Asia Minor ("Pontus, Cappadocia, Aivali: In search of Asia Minor Greek"-AmiGre) In fact, the project constitutes the first attempt to describe dialectal phenomena at a phonological, morphological, and structural level. Furthermore, it also constitutes the first attempt in Greece to combine Informatics and Theoretical Linguistics in order to facilitate the above-mentioned task. The aim here is to provide the design principles of the computational component of the project namely, an electronic dictionary and a multimedia database which would provide an innovative mechanism of storing, processing and retrieving oral and written dialectal data.

© 2014 Elsevier Ltd. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Selection and peer-review under responsibility of the 3rd International Conference on Integrated Information.

Keywords: Computational Dialectology; Asia Minor Greek; Modern Greek Dialects; Electronic Dictionaries; Multimedia Databases

1. Introduction

In recent years, research in dialectal change and language contact has come to interesting conclusions especially in the case of Germanic and Romance languages (Thomason, 2001; Matras, 2009 ; Stolz, Bakker, & Palomo (eds.), 2008). Asia Minor Greek dialects constitute a particularly interesting case in the scientific fields of dialectology and contact linguistics; although they genetically share a common Indo-European origin (Greek), they have diverged from one another partly under the influence of an Altaic language (Turkish) to such an extent that they constitute different dialects.

* Corresponding author. Tel.: +30-210-538-5824; fax: +30-210-591-0975.

E-mail address: egali@teiath.gr

It is to be noted that Greek and Turkish not only belong to different language families but to different typological groups as well (fusional vs. agglutinative). Of significant importance for the documentation of Asia Minor Greek is the contribution of the Centre for Asia Minor Studies with archival and bibliographic material (Giannakopoulos, 2003). Yet, little interest has been shown in the dialects in question with the exception of certain mentions to Cappadocian such as in Thomason (2001) and Thomason & Terrence Kaufman (1988). Therefore, an analysis of Asia Minor Greek dialects would give useful insights as for the nature and mechanism of language change within the domain of dialectal variation. On another matter, the availability of dialectal data on electronic media and the development of computational tools has greatly contributed to the advancement of research in dialectology; such is the case of the Dynamic Syntactic Atlas of Dutch Dialects (DynaSAND) (Barbiers et al., 2006) an on-line tool for dialect syntax research. It consists of a database, a search engine, a cartographic component and a bibliography concerning syntactic variation found in varieties found in the Netherlands, Belgium and France. A convincing argumentation in favour of Computational Linguistics techniques in Dialectology is reported in Nerbonne, J. (2003) while in Nerbonne, J. and Kleiweg, P. (2003) the treatment of lexical variation in LAMSAS (Linguistic Atlas of the Middle and South Atlantic States) is presented. As far as Greek dialects are concerned, no results of a computational processing of dialects are reported so far, with the exception of the electronic dictionary of Cypriot Greek (Themistocleous et al., 2012). Therefore, a computational approach to the problem is a challenge, which would contribute to the development of innovative mechanisms of processing dialectal data. Up to now, a number of linguistic studies have focused on the collection and analysis of dialectal data. Yet, none had attempted to represent both raw and processed material in the digital space. The first attempt in Greece to combine Theoretical Linguistics and Informatics for a scientific presentation of dialectal data to the academia is the THALIS program “Pontus, Cappadocia, Aivali: In search of Asia Minor Greek” (AmiGre) which aims at:

- providing a systematic and comprehensive study of Pontic, Cappadocian and Aivaliot, three Greek dialects of Asia Minor of common origin and of parallel evolution that are faced with the threat of extinction;
- digitizing, archiving and processing a wide range of oral and written data thus contributing to the sustainability and awareness of this longwinded cultural heritage.

Computational activities comprise:

- The design and development of a multimedia tri-dialectal dictionary of three Greek dialects of Asia Minor (Pontic, Cappadocian, Aivaliot). The dictionary will contain lemmata from three dialects in a comparative way.
- The design and development of a multimedia database for the archiving and processing of oral and written dialectal data
- The construction of the web site of the project where the aims, the progress and the final results will be published. (amigre.cs.teiath.gr).

In this paper, we present the design principles and the current state of development of the computational component of the project. In section 2, we present the design and implementation of the multimedia 3-dialectal dictionary. In section 3, we discuss issues concerning the design of the multimedia database. Finally, in section 4 we draw the necessary conclusions and point to future work.

Download English Version:

<https://daneshyari.com/en/article/1114781>

Download Persian Version:

<https://daneshyari.com/article/1114781>

[Daneshyari.com](https://daneshyari.com)