
Cypriot Greek Lexicography: A Reverse Dictionary of Cypriot Greek

Charalambos Themistocleous, Marianna Katsoyannou, Spyros Armosti & Kyriaki Christodoulou

Keywords: reverse dictionary, Cypriot Greek, orthographic variation, orthography standardisation, dialectal lexicography.

Abstract

This article explores the theoretical issues of producing a dialectal reverse dictionary of Cypriot Greek, the collection of data, the principles for selecting the lemmas among various candidates of word types, their orthographic representation, and the choices that were made for writing a variety without a standardized orthography.

1. Introduction

Cypriot Greek (henceforth CG) is a variety of Greek spoken by almost a million people in the Republic of Cyprus. CG differs from Standard Modern Greek (henceforth SMG) with regard to its phonetics, phonology, morphology, syntax, and even pragmatics (Goutsos and Karyolemou 2004; Papapavlou and Pavlou 1998; Papapavlou 2005; Tsiplakou 2004; Katsoyannou et al. 2006; Tsiplakou 2007; Arvaniti 2002; Terkourafi 2003).

The study of the vocabulary of CG was one of the research goals of ‘Syntychies’, a research project for the production of lexicographic resources undertaken by the Department of Byzantine and Modern Greek Studies of the University of Cyprus between 2006 and 2010. Another research goal of the project has been the study of the written representation of the dialect—since there exists no standardized CG orthography.

The applied part of the Syntychies project includes the creation of a lexicographic database suitable for the production of dialectal dictionaries of CG, such as a reverse dictionary of CG (henceforth RDCG), which is currently under publication. The lexicographic database is hosted in a dedicated webpage, which allows online searching of the database. Incorporated in the online interface of the database is a text-to-speech system, which allows the user to listen to the pronunciation of the selected lemmas. Thus, ‘Syntychies’ provides the first online version of a CG lexicographic database and a speech synthesis system for CG.

2. The RDCG

The RDCG is a lexicographic work, whose lemmas are sorted alphabetically from right to left—namely from the end of the word—and not from the beginning of the word (as is the case with standard dictionaries). Even though Greek language exhibits a long tradition of reverse dictionaries counting more than ten such dictionaries (see Kourmoulis 2002: vii-xiii), there are no reverse dictionaries for any Greek dialect. The ‘Syntychies’ project set out to fill this gap by creating the first reverse dialectal dictionary of a Greek variety.

The RDCG is an important tool for researchers studying the morphological system of CG, variety with very rich morphology. It allows researchers to examine the morphological suffixes and to make decisions about the parsing of words in various declinations and genres; one cannot make easy generalizations about the word endings by using a dictionary that is sorted left to right.

The assemble, study, and publication of the RDCG lexicographic corpus, was a very demanding and arduous task that involved many theoretical decisions, including decisions about the choices of words to be included in the database, the choice of word variants that should be treated as the main lemmas, and the orthographic representation of CG words.

3. RDCG word selection

CG is a non-standardized variety of Modern Greek and as such it includes many different pronunciations and word forms that construe geographical isoglosses (cf. Newton 1972). The choice of words to include is related to a decision about the standard form of the dialect and the vocabulary that better suits this form. The existing traditional dictionaries of CG include mostly rare, basilectal, or obsolete forms of the dialect and they are in fact a picture of CG rural communities before the urbanization that was initiated after the 1950s (cf. Papagelou 2001; Chatziioannou 1996). The RDCG aimed to include the CG vocabulary that better depicts the current state of the dialect, and as such, rare or extinct words referred in older dictionaries or word list, were excluded from the database. Word selection was based on word usage, something that was extensively examined by using questionnaires, speaker intuitions, field studies, and corpora—oral or written. The written corpus had been assembled from various sources including (a) dictionaries, (b) word lists, (c) texts from blogs, journals, and newspapers, and (d) original research studies of the oral speech. The database is an open-ended project and it is continuously updated in order to represent the current state of CG variety; and as such, an enormous task undertaken within the project was a sociolinguistic description of the CG lexicon.

4. CG orthography and written conventions

For the publication of the RDCG, the ‘Syntychies’ research project had to make hard decisions about the orthography of the dialect, since a lexicographic work must involve a certain way of normalising the spelling of words and devising rules about the written representation of the variety. Specifically, following a long tradition of using the Standard Greek orthography with some modifications for representing CG pronunciations (Chatziioannou 1996, Yiangoullis 2009, Coutsougera & Georgiou, to appear), the ‘Syntychies’ project devised an orthographic system that better represents Cypriot Greek. The proposed orthographic conventions of the RDCG are based on extensive research undertaken by the ‘Syntychies’ project (see Armosti et al., forthcoming; Armosti et al., 2011; Katsoyannou et al., 2011). The more problematic areas for writing CG are summarized in the following:

- (a) *The representation of geminates.* The phonetic repertoire of CG contains geminate sounds, which are absent from the SMG phonetic apparatus. Even though geminates are traditionally represented in SMG orthography (as a reminiscent of older forms of the language), orthographic reforms that followed the establishment of Demotic Greek (in the place of Katharevousa) as the standard language of Greece, made selective choices regarding the representation of geminates in SMG orthography in an effort to simplify the orthography (Babiniotis 2000). These choices come in stark contrast with CG geminate pronunciation. As a result, the orthographic system proposed for the RDCG, reintroduced in the orthography the writing of geminates in all words containing these consonants.

- (b) *Post-alveolar fricatives and affricates.* The phonetic apparatus of CG includes post-alveolar fricatives and affricates that do not exist in SMG. These are represented in the lexical database by means of a combining caron <ˇ> over graphemes that otherwise represent alveolar consonants. Thus, post-alveolar consonants are represented as follows: (i) the voiceless fricative [ʃ] as <ǫ̃> when singleton and as <ǫ̃ǫ̃> when geminate: e.g. <ǫ̃έριν> [ˈʃɛrin] (hand), <ǫ̃σίζω> [ˈʃ:iz:ɔ] (I tear); (ii) the voiced fricative [ʒ]: <ρύζια> [ˈri:z:ɐ] (rice-NEUT); (iii) the voiceless affricate [tʃ] as <τ̃ζ> when singleton and as <τ̃ζ̃> when geminate: e.g. <τ̃ζείνος> [ˈtʃinɔs] (that-MASC), <κουτ̃όιν> [kuˈtʃh:in] (broad bean-NEUT); (iv) the voiced affricate [dʒ] as <τ̃ζ̃>: e.g. <αντ̃ζειόν> [ɐˈndʒon] (jar-NEUT). Thus, the superscription of the combining caron over the Greek graphemes for alveolar consonants, namely <σ>, <ς>¹, and <ζ>, produces the post-alveolar fricatives (<ǫ̃>, <ζ̃>, <ξ̃>) and affricates (<τ̃ǫ̃>, <τ̃ζ̃>, <τ̃ξ̃>); moreover, the use of the combining caron differentiates between graphemes representing clusters containing alveolar fricatives (i.e. <ξ> [ks] and <ψ> [ps]) and graphemes representing clusters containing post-alveolar fricatives (i.e. <ξ̃> [kʃ] and <ψ̃> [pʃ]).
- (c) *Final /n/.* Many words, such as neuter nouns and adjectives, end in /n/ in CG, but not in SMG: for example the word for ‘school’ is written as <σχολείο> in the SMG orthography, but as <σχολείον> in the lexical database, wherein word-final /n/ is always represented.
- (d) Synizesis is not represented in the lexical database since it is in most cases predictable from CG phonology—however other dictionaries represent it by using various symbols (see Yiangoullis 2011).

5. Future Research

RDCG is an innovative, targeted dictionary, as not only is the first published reverse dictionary of CG, but is also the first reverse dictionary of any Modern Greek dialect. CG lexicography lacks modern, scientific approaches, a gap which the ‘Syntychies’ research program sets out to fill by making use of modern methodologies—something that resulted in the RDCG. ‘Syntychies’ aspire to provide a standard lexicological reference material for CG. Furthermore, ‘Syntychies’ aim to provide further lexicographic recourses not only to academics, researchers, and scholars, but to anyone interested in CG. Since the lexical database of ‘Syntychies’ is open-ended and always updating, RDCG provides a current picture of CG variety available to anyone. There are also further plans for expanding this research and for providing target-specific dictionaries such as the RDCG.

Note

¹ The <ς> variant is an allograph of the <σ> grapheme when in word-final position.

References

A. Dictionaries

- Chatziioannou, K. 1996.** *Ετυμολογικό Λεξικό της Ομιλουμένης Κυπριακής Διαλέκτου: Ιστορία Ερμηνεία και Φωνητική των Λέξεων με Τοπωνυμικό Παράρτημα* [An Etymological Dictionary of Presently Spoken Cypriot Greek: History, Explanation and Phonetics of Words; With an Appendix of Toponyms]. Nicosia: Tamassos.

- Kourmoulis G. 2002.** *Αντίστροφον λεξικόν της Νέας Ελληνικής* [Reverse dictionary of Modern Greek]. (2nd edition.) Athens: Papadimas.
- Papagelou, R. 2001.** *Το Κυπριακό Ιδίωμα: Μέγα Κυπρο-Ελληνο-Αγγλικό (και με Λατινική Ορολογία) Λεξικό: Ερμηνευτικό-Ετυμολογικό-Προφοράς-Ορθής Γραφής* [The Cypriot Greek Dialect: Great Cypriot - Standard Greek - English (And Latin) Dictionary of Etymology - Pronunciation - Correct Spelling]. Athens: Iolkos.
- Yiangoullis, C. 2009.** *Θησαυρός Κυπριακής Διαλέκτου* [Thesaurus of the Cypriot Greek dialect]. (3rd edition.) Nicosia: Theopress.

B. Other literature

- Armosti, S., K. Christodoulou, M. Katsoyannou, and C. Themistocleous forthcoming.** ‘Writing in Cypriot Greek: The Need for Standardisation and its Importance for Dialectal Lexicography.’ In *Dialogue on Dialect Standardization*. Cambridge: Cambridge Scholars Publishing.
- Armosti, S., M. Katsoyannou, K. Christodoulou, and Ch. Themistocleous to appear.** ‘Τάσεις των Κυπρίων ως προς τη γραπτή απόδοση των μεταφατνιακών συμφώνων της κυπριακής. [Tendencies of Greek Cypriots in writing Cypriot Greek post-alveolar consonants]’ In *Proceedings of the 10th International Conference of Greek Linguistics*.
- Arvaniti, A. 2002.** ‘Διμορφία, διγλωσσία και η εμφάνιση της Κυπριακής κοινής [Diglossia, bilingualism and the emergence of the Cypriot Koine].’ In *Recherches en linguistique grecque*, vol. I, 75–78. Paris: L’Harmattan.
- Babiniotis, G. 2000.** *Ιστορία της Ελληνικής Γλώσσας. Με Εισαγωγή στην Ιστορικοσυγκριτική Γλωσσολογία* [History of the Greek Language. With Introduction in Historical Comparative Linguistics]. Athens: Romanos.
- Coutsougera, P. and G. V. Georgiou. to appear.** ‘An Orthographic System for Cypriot Greek.’ In *Proceedings of the XXX Colloque International de Linguistique Fonctionnelle*. Paris: L’Harmattan.
- Goutsos, D., and M. Karyolemou 2004.** ‘The Sociolinguistics of Cyprus I: Studies from the Greek Sphere.’ *International Journal of the Sociology of Language* 168: 1–17.
- Jurafski, D. and J. Martin. 2000.** *Speech and Language Processing*. New Jersey: Prentice Hall.
- Katsoyannou, M., A. Papapavlou, P. Pavlou and S. Tsiplakou. 2006.** ‘Διδιαλεκτικές Κοινότητες και Γλωσσικό Συνεχές: η Περίπτωση της Κυπριακής [Bidialectal Communities and the Linguistic Continuum: the Case of Cypriot].’ In *Proceedings of the 2nd International Conference on Modern Greek Dialects and Linguistic Theory*, 156–171.
- Katsoyannou, M., S. Armosti, K. Christoloulou, and C. Themistocleous. 2011.** Σύγχρονη Κυπριακή Λεξικογραφία: η Ηλεκτρονική Βάση Δεδομένων «Συντυχιές» [Contemporary Cypriot Lexicography: the 'Syntychie's' Electronic Database]. Paper presented at the 9th International Conference on Greek Research, Adelaide, South Australia, 30 June – 3 July 2011.
- Newton, B. 1972.** *Cypriot Greek. Its phonology and inflexion*. The Hague: Mouton.
- Papapavlou, A. 2005.** *Contemporary sociolinguistic issues in Cyprus*. Thessaloniki: University Studio Press.
- Papapavlou, A. N., and P. Pavlou 1998.** ‘A Review of the Sociolinguistic Aspects of the Greek Cypriot Dialect.’ *Journal of Multilingual and Multicultural Development* 19.3: 212–220.
- Terkourafi, M. 2003.** ‘The Cypriot koine: A recent development?’ In *Sixth International Conference of Greek Linguistics, Rethymnon, 18-21 September 2003*, Rethymnon:

Linguistics Lab, 1-26. Retrieved from

<http://www.philology.uoc.gr/conferences/6thICGL/ebook/g/terkourafi.pdf>

Tsiplakou, S. 2004. ‘Στάσεις Απέναντι στη Γλώσσα και Γλωσσική Αλλαγή: μια Αμφίδρομη Σχέση; [Attitudes towards Language and Language Change: a Two-Way Relationship?].’ In *Proceedings of the 6th International Conference on Greek Linguistics*. Retrieved from <http://www.philology.uoc.gr/conferences/6thICGL/>

Tsiplakou, S. 2007. ‘Γλωσσική εναλλαγή, ανταγωνιστικές γραμματικές και διαγλώσσα: Ακόμα μια διάσταση του Κυπριακού. [Code switching, competitive grammars and interlanguage: Another dimension of the Cypriot problem].’ In *Proceedings of the 8th International Conference on Greek Linguistics*: 1195–1209.