THE HISTORICAL DICTIONARY OF MODERN GREEK:
DIALECTOLOGICAL ISSUES

Christina BASSEA-BEZANTAKOU & Io MANOLESSOU
Academy of Athens
chbassea@academyofathens.gr / manolessou@academyofathens.gr

Abstract
This paper aims to give a general presentation of the Historical Dictionary of Modern Greek (ILNE): aims, history, and recent innovations, with focus on specific problems deriving from the updating of its corpus and methodology. Emphasis is given to the issue of the source language, which, in contrast to most historical dictionary projects world-wide, is not only the modern standard language but also the dialects. The necessity of inclusion of dialectal data in a historical dictionary of the standard language is discussed and illustrated at length, while other issues involving the treatment of dialectal data, such as dating and phonetic transcription, are also touched upon.

Keywords
historical lexicography, dialect lexicography, Greek, diglossia, phonetic transcription

Resumen
El propósito principal de este artículo es ofrecer una presentación general del “Diccionario Histórico del Griego Moderno” (ILNE): sus objetivos, su historia y sus innovaciones recientes, dedicando una especial atención a los problemas específicos derivados de la actualización de su corpus y de la metodología. Se hace hincapié en la cuestión de la lengua de origen, la cual, a diferencia de la mayoría de proyectos internacionales de diccionarios históricos, no sólo es la lengua estándar moderna, sino también la dialectal. Se discute ampliamente y con detalle sobre la necesidad de incluir datos dialectales en un diccionario histórico de la lengua estándar, aunque otros problemas relacionados con
el tratamiento de los datos dialectales, como las citas y la transcripción fonética, también están considerados.

**Palabras clave**

lexicografía histórica, lexicografía dialectal, griego, diglosia, transcripción fonética

---

1. **Introduction-Presentation of the *Historical Dictionary of Modern Greek* (ILNE)**

The “*Historical Dictionary of Modern Greek*” (*Istrikon Lexikon tis Neas Ellinikis*-ILNE) is the national dictionary enterprise of Greece, bearing all the typical features of such works: historical perspective, detailed coverage, full etymological treatment, variant forms and their distribution, rich documentation with examples from spoken language and from the literary canon, folklore information (folksongs, proverbs, curses), use of words as proper names and place-names, bibliography, etc. The time-span covered by the ILNE is 1800-today, with a retrospective examination up to the first appearance of meanings and forms.¹ Despite the recent growth of interest in historical lexicography, the ILNE is never discussed in recent major works offering overviews of international historical lexicographical practice (e.g., Merkin 1983, Considine 2010). The aim of this paper is to fill this gap, adding an interesting example to the metalexicographic discussion on historical and dialectal dictionaries.

The ILNE was established in 1908, on the initiative of Georgios Chatzidakis, the founder of the linguistic science in Greece, and was one of the first research centers to be placed under the auspices and superintendence of the Academy of Athens (1927).² Its initial purpose was the compilation of a single dictionary of all the Greek language, from antiquity until modern times, which would be “historical” because it would provide information on the history of words, i.e. their phonological, morphological and

---

¹ For the features of a typical historical dictionary see Reichmann (1990, 2012); according to his typology, the ILNE is a historical dictionary of the present language (*gegenwartsbezogen*), with a diachronic perspective. For more details on the ILNE as a historical dictionary see Manolessou (2012b). For a shorter, English-language presentation of the ILNE, see Bassea-Bezantakou (2010).

² For the history of the ILNE see Bassea-Bezantakou (2010), Giakoumaki, Karantz & Manolessou (2004), Charalambakis (2003).
semantic evolution along the time axis. As its prologue claimed, it would constitute “proof of the linguistic unity of the nation through the centuries, and a monument to the immortality of the Greek race” (ILNE, Prologue, vol. I (1933: η΄)).

This over-ambitious initial purpose, understandable within the framework of late 19th century nationalism and romanticism, was of course unachievable with the means available at the time. Thus, the scope of the dictionary was soon revised and reduced to the spoken Modern Greek language, “both the commonly spoken language and its dialects”, and the chronological limits of the dictionary’s coverage were set from 1800 until today (today being 1935, the date of the publication of the Historical Dictionary’s Manual of Regulations). Since then, the rate of progress has been rather slow: five volumes have appeared, covering the letters A-D, the last published in 1989. The current volume (the 6th) is due to appear in 2014. However, during the last 5 years, a major revision programme has been taking place at the ILNE, described in the revised Manual of Regulations (2012), and resulting in important changes in methodology, data processing and data presentation, which bring the ILNE up to date with modern historical lexicographical practice.

2. Lexicographical issues I: Data coverage

As pointed out above, the ILNE constitutes in most respects a historical dictionary: it is a grand-scale and long-term multi-volume enterprise, which aims to document the national language in all its detail and in its diachronic dimension. In this respect, it resembles projects like the DWB, the OED or the WNT. In one major respect, however, the ILNE differs from such enterprises: it aims to include not only the standard national language, but also the dialects. This objective multiplies

---

3 For the patriotic and romantic perspective of 19th c. historical lexicography, of which the ILNE is a typical instance, see Reichmann (2007), Considine (1998) and especially Reichmann (2012: ch. A.5).

4 See ILNE-MR (2012) and the special issue of the the periodical Lexicographicon Deftion 26 (2012), which contains several metalexicographical studies discussing various aspects of the new MR.

5 Given the “national” perspective of the major historical lexicographic enterprises, the absence of dialectal material is not surprising. However, the recent tendency in historical lexicography is to include much more dialectal data, as their great contribution to the understanding of language history is
considerably the size of the material to be treated and consequently lengthens, even more considerably, the time required for the completion of the project. Nevertheless, it was and remains a necessary decision.

At the time of the Dictionary’s inception, the Greek language was in a state of diglossia; it was in fact the paradigm case of diglossia, considering standard references to the subject such as Ferguson (1959). This means that the “national language” during the 19th and most of the 20th century was a high register, learned and archaic, and to a certain extent “artificial”, language at all levels of analysis: phonology, morphology, syntax and vocabulary. The very conservative orthography of Modern Greek (MG) helped to mask the actual phonetic distance between Ancient (AG) and puristic forms. To give a concrete example, the following Table 1 gives the AG and MG graphematic and phonetic forms of three very common words:

<table>
<thead>
<tr>
<th></th>
<th>AG</th>
<th>MG-puristic</th>
<th>MG-vernacular</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘king’</td>
<td>βασιλεύς</td>
<td>βασιλεύς</td>
<td>βασιλιάς (βασιλέας, βασιλιός)</td>
</tr>
<tr>
<td></td>
<td>basileú̱s</td>
<td>vasiˈlefs</td>
<td>vasiˈlas (vasiˈleas, vasiˈlós)</td>
</tr>
<tr>
<td>‘write’ 3rd. pl. pass. imperf.</td>
<td>ἐγράφοντο</td>
<td>ἐγράφοντο</td>
<td>γράφονταν, γραφόντανε, γραφόντουσαν</td>
</tr>
<tr>
<td></td>
<td>eˈgrafondo</td>
<td>γραφόνταν</td>
<td>γραφόντανι, γραφόντανη, γραφόντουσαν</td>
</tr>
<tr>
<td></td>
<td>ἐγράφοντο</td>
<td>eˈgrafondo</td>
<td>γραφόνταν, γραφόντανε, γραφόντουσαν</td>
</tr>
<tr>
<td>‘fish’</td>
<td>ἰχθύς</td>
<td>ἰχθύς, ὄψαριον</td>
<td>ὄψαρι</td>
</tr>
<tr>
<td></td>
<td>ἰχθύς</td>
<td>ἱχθύς, ὀψάριον</td>
<td>ὄψαρι</td>
</tr>
<tr>
<td></td>
<td>ἰχθύς</td>
<td>ἱχθύς, ὀψάριον</td>
<td>ὄψαρι</td>
</tr>
<tr>
<td></td>
<td>ἱχθύς</td>
<td>ἱχθύς, ὀψάριον</td>
<td>ὄψαρι</td>
</tr>
</tbody>
</table>

Table 1. The words ‘king’, ‘write’ and ‘fish’ in Ancient and Modern Greek

Therefore, if one really wished to investigate the lexical history of Greek, one simply had to turn to the dialects and to the “language of the simple people”, which reflected linguistic history more directly and faithfully. But even after the resolution of the Greek “language question”, within the 20th century, in favour of the so-called gradually being recognized. On this topic see Kühn & Puschel (1983), Durkin (2009) and the discussion in Manolessou (2012b: 30-35).

On the phenomenon of diglossia in Greek and the two varieties (usually termed “katharevousa”, the high purist variety, as opposed to the vernacular “demotic”), see Browning (1982), Mackridge (1990, 2009); see also Tseronis (2002) for its impact on MG lexicography.
“demotic” language, it remains impossible to do serious diachronic work on Greek without reliance on dialectal data. The centuries-long diglossia means that:

1) Written sources consciously avoid the recording of “vernacular” forms and instead give archaic variants which hide linguistic evolution. This hinders the dating of changes, as well as the investigation of their gradual spread. It is frequently the case that written evidence exists only for the initial, original form in Ancient Greek, and the final outcome in MG, whereas the intermediate steps or stages, which must have existed in Medieval or Early Modern times are unattested, and recoverable only through the oral evidence of conservative dialects which preserve medieval characteristics. Conversely, many words and phenomena well attested in Medieval texts are absent from Standard MG but regular in most or all modern dialects. For example, the active perfect tense is expressed in Standard MG through the auxiliary ‘have’ + infinitive (1a). This formation is very sporadically attested in late medieval texts from the 15th century onwards, and becomes incorporated in the grammatical descriptions of MG only in the 19th century. Most of the dialects still show the regular medieval formation, which is ‘have’ + passive participle (1b):  

1a. ἔχω γράψει ['exo ɣrapsi] ‘I have written’

   have.1.sg write.infin.

1b. ἔχω γραμμένο ['exo ɣrammeno] ‘I have written’

   have.1.sg. write.perf.pass.ptcpl.n.acc.sg.

2) Phonetic changes usually apply irregularly and sporadically in Standard Modern Greek, but with the expected exceptionless (“lautgesetzlich”) regularity in the dialects. Therefore, the historical phonology and morphology of MG are better investigated through dialectal evidence. A characteristic example is the /l/ > /r/ change in Medieval and Modern Greek. Since the Hellenistic period, Greek phonotactics

---

7 This was a basic principle of Georgios Chatzidakis, the founder of the ILNE, cf. his classic works Hatzidakis (1892) and Chatzidakis (1905-1907).
8 On the appearance of the ‘have + infinitive’ perfect in MG and MG grammars see Manolessou (2012a); on the variants of the perfect in MG dialects see Ralli, Melissaropoulou & Tsolakidis (2007).
9 The phenomenon is discussed in detail in Manolessou & Toufexis (2009).
include a constraint which disallows clusters of lateral /l/ + obstruent, changing the lateral /l/ to /r/, a phenomenon known in the literature as “liquid interchange”. The change appears in a very low number of lexical items in Standard Modern Greek (2a), most of them retaining the “original” form (2b) and thus this phonetic change should be characterized as “sporadic”, as far as Standard Modern Greek is concerned. However, if one turns to dialect data, not only will one find the expected regular outcomes of the change in inherited words (3a), but one will see the change affecting even relatively recent loanwords (3b):

2a. AG ἀδελφός [adelpʰós] > MG αδερφός [aderˈfos] ‘brother’
   AG ἀλμυρός [halmyrós] > MG αρμυρός [armiˈros] ‘salty’

2b. AG ἐλπίς [elpís] > MG ελπιδά [elˈpiða] ‘hope’
   AG τόλμη [tólme:] > MG τόλμη [ˈtolmi] ‘courage’

3a. MG-dialectal ερπίδα [εɾˈpiða], τόρμη [ˈtormi] (Cyprus)

3b. Engl. penalty (football term) > Cyprus πέναρτι [ˈpenarti]
   Germ. Waltz > Naxos δάροι [ˈvarsi]
   Russ. Bolshevik > Naxos μπορσεβίκος [borseˈvikos]

3) The Modern Greek vocabulary contains a very high number of etymological doublets, i.e. pairs of words with the same etymon, one of which has undergone regular change while the other has remained relatively unchanged. In examples (4) and (5), the words in question have two forms in Standard MG: one has undergone the phenomenon of synizesis (semivocalisation of [i] when followed by a vowel) and one has not, the ‘older’ form retaining an original meaning and the innovative form showing a meaning change; in some conservative MG dialects which have not undergone the process of synizesis, the original form has the new meaning:

4a. AG χωρίον [kʰɔːrion] ‘inhabited area, textual passage’
4b. MG χωρίο [xoˈrio] ‘textual passage’ ~ χωριό [xorˈjo] ‘village’
4c. MG-dialectal χωρίο [xoˈrio] ‘village’ (Pontus, Mani)

5a. AG δουλεία [duːˈlēia] ‘slavery’
5b. MG δουλεία [ðuˈlia] ‘slavery’ ~ δουλειά [ðuˈʎa] ‘work, job’
5c. MG-dialectal δουλεία [ðuˈlia] ‘work, job’

4) A large part of the Standard Modern Greek vocabulary are recent acquisitions or re-acquisitions that have entered the language from the archaic, learned register during the 19th and early 20th centuries. These lexical items are the results of a conscious ‘purification’ of the language (hence the term ‘puristic’) which came to replace very common vernacular vocabulary of foreign origin (mainly Italian and Turkish). This vocabulary largely survives in the dialects but not in the standard. In example (6), the original names of professions are of Italian and Turkish origin respectively; they have been replaced in the Standard by ancient reintroductions or new learned coinages and survive only in dialects or as proper (last) names.

6a. MG φαρμακοποιός [farmakopiˈos] ‘chemist’ (learned purist coinage)
   MG-dialectal σπετζέρης [speˈtseris] ‘chemist’ < Ital. speziere

6b. MG κηπουρός [kipuˈros] ‘gardener’ (< AG κηπουρός)
   MG-dialectal μπαξεβάνης [bakseˈvanis] ‘gardener’ (< turk. bahçivan)

More generally, the MG dialectal vocabulary presents a high number of archaic retentions at all levels of lexical analysis (phonology, morphology, semantics), which differ from the “archaicity” of the Standard language: they are the result of direct uninterrupted transmission, generation to generation, through the spoken language; on the contrary, the archaic and conservative lexical features of Standard MG are mostly the result of its diglossic history, i.e. they are “learned” internal loans from the purist language or directly from Ancient Greek. For this reason, the term “archaism” has a special meaning in Modern Greek dialectology, which is different from that attributed to it in usual linguistic practice.

Linguistics dictionaries \(^{10}\) define archaism as an obsolete feature surviving marginally in the standard language (in poetry, nursery rhymes, place-names etc.),

\(^{10}\) See Bussmann (1996: s.v.), Crystal (2008: s.v.), Matthews (2007: s.v.). See also Cherubim (1988) for archaisms in the German vocabulary.
patently belonging to an earlier stage of linguistic history, and having connotations of obsolescence, antiquity, or elevation. In Modern Greek dialectology, archaisms are normal every-day features of dialectal language, which present conservative characteristics absent from Standard Modern Greek, and which are not due to the influence of the written, learned tradition and the purist language (cf. the classic work by N. Andriotis (1974), *Lexikon der Archaismen in neugriechischen Dialekten*). Such dialectal archaisms were accorded special importance in past decades, as they were considered incontestable proofs of the historical continuity of Greek and the strong connections of the modern language to its ancestor form, Ancient Greek, with its unsurpassed international cultural prestige. Dialectal archaisms are still important both in Modern Greek linguistic history (as they aid dialectal classification and the tracing of phonetic, morphological and semantic isoglosses and shed light on the causes inhibiting language change) and in Ancient Greek linguistic and literary studies, as they assist in the better understanding of the Ancient Greek vocabulary, corroborating a dubious meaning or illuminating unknown words.

As already mentioned, dialectal archaisms may be further subdivided in phonetic archaisms (7), semantic archaisms (8) and lexical archaisms (9), as can be seen from the examples below:

7a. AG ζημία [dzɛ:míα] > MG ζημιά [ziˈmŋa] ‘loss, damage’
    Pontic ζεμία [zeˈmia]

7b. AG βήξ, acc. βῆχα [bɛ:ˈks], [bɛːˈkʰa] > MG βήχας [ˈvixas] ‘cough’
    Pontic βέχας [ˈvexas]

7c. AG ἐργάτης [ergátɛs] > MG ἐργάτης [erˈɣatis] ‘worker’
    Pontic ἐργάτες [erˈγates]

The AG (Attic) phoneme /ɛ:/ evolved in Koine times to /i/, and this outcome appears in Standard MG and all MG dialects (Horrocks 2010: 160-163). Pontic is the only exception: in many words (but not with absolute regularity) the AG vowel quality
was preserved, and thus Pontic retains /e/ where Standard MG and all other varieties have /i/ in both stems (7a, b) and suffixes (7c).\textsuperscript{11}

8a. AG δέμα [déma] ‘band, rope’ > MG δέμα [ˈðema] ‘bundle, pack’
   MG-dialectal δέμα [ˈðema] ‘band, rope, cord’

8b. AG πάγος [págos] ‘crag, hill, coagulated liquid’ > MG πάγος [ˈpágos] ‘ice’
   Peloponnesian πάγος [ˈpágos] ‘rock’

8c. AG λούομαι [lu:ˈomai] ‘I wash myself’ > MG λούζομαι [ˈluzome] ‘I wash my hair’
   Cypriot λούν-νουμαι [ˈlun:ume] ‘I wash myself’

8d. AG ἀσήμαντος [aˈsimandos] ‘unmarked, not branded’ > MG ασήμαντος [aˈsimandos] ‘insignificant, unimportant’
   MG-dialectal ασήμαντος [aˈsimandos] ‘unmarked, not branded’ (of sheep)

In all four cases (8a, b, c, d) the lexical item is present both in Standard MG and its dialects; but while in Standard MG the word has undergone a semantic change and its original sense has been lost, in the dialects the original meaning is retained in normal use, sometimes in parallel with the Standard meaning and sometimes exclusively.

9a. AG ἀζα [ádza] ‘heat, dryness, dirt’
   MG-dialectal ἀζα [aza] ‘cinder’ (Euboea, Skyros), ‘dust’ (Peloponnesian)

9b. AG δεῖσα [déisa] ‘slime, filth, mud’
   MG-dialectal δεῖσα [ðisa] ‘heavy mist’ (Pontic)

9c. AG σῆτες [sέ:tes], Doric σάτες [sá:tes] ‘this year’ (adverb)
   MG-dialectal σάτσ [ˈsatsʰ] ‘this year’ (Tsakonian)

In all three cases the AG lexical items are rare words with very few attestations in the extant corpus of AG literature (see LSJ s.vv.). They are also absent from Standard MG. However, they are retained in some MG dialects, either in their original meaning (9c) or with slight semantic evolution (9a, 9b).

\textsuperscript{11} For a detailed description of the phenomenon and its diachronic attestations see Manolessou & Pantelidis (2011).
As a further linguistic argument concerning the inclusion of dialectal data in the ILNE, it should be considered that the time-span covered by the dictionary starts, as already mentioned, from the beginning of the 19th century, i.e. from 1800. A considerable portion of the vocabulary which was “common vernacular” during the 19th century has, however, become obsolete in the 20th, especially in domains such as agriculture, industry, professions, instruments, customs etc. This vocabulary has been retained in many dialects, where antiquated practices or customs still survive. Therefore, the study of modern dialectal vocabulary is absolutely necessary for the investigation of 19th century language and, more generally, 19th century extralinguistic reality (objects, activities, customs).\textsuperscript{12}

Two final, ideological, causes lie behind the inclusion of dialectal data in the “national” Greek dictionary, in contrast to the practice of its contemporary models such as the DWB and the OED: one of the main socio-cultural and ideological purposes served by 19th century historical lexicography was the construction of a “national identity” (Reichmann 2012). This was achievable firstly through the promotion of a standard linguistic form, as seen in the “great”, “mainstream” literary authors and as cultivated by state education and administration;\textsuperscript{13} such a model could not be served by the exhibition of the bewildering variation of phonetic and morphological variants or the fluctuation of meanings presented by the “dialects”, to which the standard language was opposed. Furthermore, the historical examination of dialectal variant forms and the tracing back in time of their etymological origin could put into question the ultimate unity of the “language” and therefore of the “nation”, since in the case of many modern languages (e.g. Italian, German), their dialectal varieties go as far back in time as “the standard” itself.

In the case of Greek, neither of these two considerations entered into play: the officially promoted language was the high purist variety, which was not included in the dictionary at all (since, as already demonstrated, it was a very poor basis for the investigation of linguistic history); therefore it could not be menaced by the inclusion

\textsuperscript{12} On this issue see Bassea-Bezantakou (2012), where the value of encyclopedic information that should be provided on the basis of dialectal vocabulary is discussed.

\textsuperscript{13} On this issue see also Beltrami & Fornara (2004), concerning the exclusive promotion of the dialect of Tuscany in Italian lexicography.
of the rich phonetic, morphological and semantic variety of dialectal forms existing alongside the common (supra-regional) “demotic” ones. It could even be argued that the exposition of this wide variety of forms and meanings indirectly and covertly served the purpose of proving that a high, purist language, above and beyond the unsystematicity of the vernacular, was actually necessary. Furthermore, the examination of the historical origin of dialectal forms posed no threat to the proof of the ultimate origin of “the language” and “the nation”: all dialects of Modern Greek have a well-traced common descent from Koine (Hellenistic) Greek, which in turns derives from a single ancient Greek dialect, Attic.14

With the resolution of the language question in 1976 in favour of the ‘demotic’ language as the official language of Greece (Mackridge 1990), the ILNE found itself, almost accidentally, in the unique for historical lexicography position of having to treat its standard language together with its dialectal varieties in the same dictionary.

### 3. Lexicographical issues II: Methodological innovations

#### 3.1 Dating

One of the most important changes in the methodology of the ILNE, as set out in the new Manual of Regulations, concerns the dating of lexical items and of senses. Of course, dating was always a concern for the ILNE as a historical dictionary, and under the old regulations it was necessary to provide a rough dating (by period, i.e. Ancient, Hellenistic, Medieval etc.) for the first appearance of each lexical item, as well for each of its senses. However, recent technological advances have truly revolutionized the diachronic investigation of Greek, and this is something that needs to be taken fully advantage of in the future volumes of the ILNE. The advances consist of a) recent completion of major historical dictionaries for earlier phases of Greek, both print and

14 For the history of the modern Greek languages see Horrocks (2010). The only exception to the ultimate Attic origin of all modern Greek dialects is Tsakonian, descended from Doric (see Pernot 1934 for a detailed description of its synchrony and diachrony).
electronic b) development of electronic historical corpora for Greek and c) mass digitization of rare early modern printed editions, which, being out of copyright, are easily accessible online.

The earliest phase of Greek,\textsuperscript{15} Mycenaean, is covered by the \textit{DMic}. Ancient Greek up to the 6th century A.D. is of course covered by the monumental \textit{LSJ}, available online. At the same time, a new major lexicographical project for Ancient Greek, the \textit{DGE} (still under way-available online) provides much additional data, mainly from papyri and inscriptions, as well as a more up-to-date semantic analysis. The Medieval period is covered by the Dictionary of Kriaras (currently up to the letter Rho- available online) and by the \textit{LBG} (currently up to Tau). Most dictionaries are comparatively recent, or very recent achievements which were unavailable for the earlier volumes of the ILNE.

The most important technological advance for the history of Greek is the development of the \textit{Thesaurus Linguae Graecae}, of the University of Irvine, California. The TLG (www.tlg.uci.edu) is a searchable electronic corpus that contains not only the totality of AG literature, but also the majority of sources of the Byzantine period. In the last five years, a large amount of sources from later periods, reaching up to the 17th century has been added. Complementing the TLG, two other searchable electronic corpora for Ancient, Koine and Early Medieval Greek are available. The Cornell Epigraphy Project (http://epigraphy.packhum.org), which offers full access to Ancient and Christian inscriptions, and papyri.info (www.papyri.info), which offers the totality of the published Greek papyri. In the last decade, two electronic corpora for Standard Modern Greek have also been developed and are available online, SEK (www.sek.edu.gr) and the Hellenic National Corpus (HNC – hnc.ilsp.gr ).\textsuperscript{16}

With the help of these resources, it is possible to trace the first appearance of a lexical item with great precision, as well as with great speed and minimal effort. It is also much easier to trace the history of senses, as the lexicographer has now at his fingertips, with a single click, not only the dictionary information for the past history of word, but the full range of attestations in the original sources, which can be examined

\textsuperscript{15} For an overview of Greek historical lexicography see Manolessou (2012), and more details in Zgusta & Georgacas (1990), Georgacas & Georgacas (1990), Kramer (2011) and Kazazis (2003).

\textsuperscript{16} The available electronic corpora for Greek are discussed in Manolessou & Toufexis (2011), Konstantinidou & Tzamali (2012).
first-hand. One might think that the lexicographer thus runs the risk to drown in too much data. Experience has shown that this danger exists mainly for “major” lexical items such as “love, take, have, show” etc., whereas less central words can be treated fully within a reasonable time-span.

The availability of new tools for the investigation of the history of Greek, touches upon another important issue, which greatly influences the time-span required for the completion of the ILNE project. In contrast to most if not all European languages, Greek has a very long, continuously attested history. There is no clear limit for the start of “Modern Greek”; contrast the relative ease with which one can date the beginning of, say, French or Italian as opposed to Latin. It is of course impossible, as already discussed in the introduction, to cover the whole history of Greek within a single dictionary, and the ILNE cannot hope to achieve that. However, when tracing back the history of modern words and senses, one will fall very short of scientific truth if one stops at some arbitrarily set point. It will also be a pity to reject the rare opportunity to observe lexical and semantic change for a length of 3,500 thousand years. But the necessity of tracing many vocabulary items so far back in time greatly increases the amount of sources that have to be examined.

Turning now to the third important technological advance which is being taken advantage of in the new and planned volumes of the ILNE: this concerns the mass digitization, by independent organisations (libraries, universities, private collectors), of old and rare publications, which are of great use for the dating of lexical items. This concerns not so much the ancient and medieval period, which is mostly covered by the new electronic corpora and dictionaries as discussed above, but rather the Early Modern period, i.e. the 16th, 17th 18th and 19th centuries. These centuries are not so well researched by modern scholars and they are not included in most of the extant historical dictionaries or corpora. Therefore, the dating of lexical items first appearing during these centuries is extremely difficult. The solution adopted by the ILNE is to exploit, as far as possible, printed dictionaries published during these periods. Until recently, most of these dictionaries were hard to locate and to use, as they existed only in rare copies in various libraries around the world. In the last decade, however (or rather in the last five years), they have become available to the scholarly public
through the internet. The ILNE has compiled a full list of such lexicographical works, and has managed to acquire digital copies of almost all of them, thus greatly facilitating and speeding up the lexicographers’ work.

3.2 Phonetic transcription: problems and solutions

The other important innovation of the ILNE’s new regulations concerns the phonetic transcription of linguistic forms. Up to now, the ILNE used only the conventional Greek alphabet and orthography, with a number of custom-made additional symbols in order to represent dialectal sounds non-existent in the standard language. This practice is well-established in all European countries, as the International Phonetic alphabet does not have a long tradition in the field of dialectology and dialectal lexicography. For example, German-language dictionaries prefer the Teuthonista system, Italian dialectology relies on the similar Böhmer-Ascoli system, while French dialectology often employs the Rousselot-Gillieron alphabet. Following this practice, the ILNE developed, already from its first volumes, a custom-made phonetic system based on Greek orthography.

However, from the next volume of the ILNE, all word-forms will also be accompanied by a phonetic transcription in the IPA. The gain from such a measure is obvious: The transcription will be transparent, consistent and standardized. Forms will be accessible to the international scholarly public, not only to Greek-speaking scholars. Of course, the decision to adopt the IPA creates major methodological and practical problems, which will be shortly reviewed here.

Firstly, the decision which symbol to use requires information which is not available for most Modern Greek dialects, as dialectal phonetics and phonology are relatively new disciplines, with few relevant publications. Therefore, the lexicographer is called upon to choose a symbol denoting specific articulatory

---

17 See ILNE-MR (2012) for a list of the relevant works, and Perakis (1994) for further details.
18 For an overview of dialectological practice with respect to phonetic transcription see Manolessou, Beis & Bassea-Bezantakou (2012), and Schiefer & Pompeiolo-Marchall (1996).
19 A brief history of phonetic research in MG dialectology is offered in Manolessou, Beis & Bassea-Bezantakou (2012).
properties, without empirical and experimental knowledge of the articulatory or acoustic properties of MG dialectal sounds.

Furthermore, older studies of MG dialects often employ an ad hoc symbolism which is no longer comprehensible, or describe phonetic phenomena with non-standard, antiquated or amateur terminology. When transcribing a word, the ILNE researcher is thus frequently forced to “guess” what the older, manuscript or print source must have meant, on the basis of more recent scientific work on the dialect in question. If that is still extant, this is to a certain degree possible, if a bit arbitrary. But if the dialect is now extinct, it becomes almost impossible. Up to a point, the problem can be overcome by recourse to the ILNE’s rich sound archive (containing several hundred hours of recorded dialect material, from the 1950’s onwards).\textsuperscript{20}

A concrete example of the problems involved in the IPA transcription of written dialectal material is offered by double consonants in MG dialects:\textsuperscript{21} Although the MG phonological system does not allow double (geminate) consonants, a number of dialects, namely the peripheral dialect groups of South Italy and of the South-East (Dodecanese, Cyprus, Chios etc.), possess more complex phonological systems which involve double (geminate) consonants, both as a survival from Ancient Greek and as an innovative evolution. The precise phonetic realization of double consonants is relatively easy to understand and to represent in the case of continuants (fricatives, nasals, liquids etc.): they are consonants with long (double) duration, and in impressionistic non-IPA transcriptions they are represented through the repetition of the consonant in question, with or without an intervening dash (10). But in the case of stops, several articulatory parameters come into play, with considerable cross-dialectal variation: duration, aspiration, different place of articulation etc. These are graphematically represented with a bewildering and inconsistent variety in the dialectal written sources and the researcher has to decide whether these graphemes are meant to represent only lengthening, lengthening and aspiration, only aspiration, only affrication, lengthening and affrication, etc. (11). Sometimes, modern fieldwork in

\textsuperscript{20} The oldest recordings of a MG dialect were made by Paul Kretschmer in 1905 (kept now in Vienna).
\textsuperscript{21} The issue is discussed in detail in Manolessou & Bassea-Bezantakou (2012). An easily accessible description of geminates in MG dialects can be found in Newton (1972: 89-93), Seiler (1958).
the area provides the answer; at other times no reliable modern work exists and one has to rely on the impressionistic description accompanying the graphematic representation.

10a. MG δάλασσα [ˈθalasa] ‘sea’ (< AG δάλασσα [tʰalas:a])

MG-dialectal δάλασ-σα [ˈθalas:a]

10b. MG νύφη [ˈnifi] ‘bride’

MG-dialectal νύφη, νύφ-φη [ˈnifi: (< AG νύμφη [nýmpʰɛ:]])

10c. MG απλώνω [aˈplono] ‘I spread’

MG-dialectal απλώννω, απλώ-νω, πλώ-νω [aˈplon:o], [ˈplon:o]

11. MG κουπί [kupi] ‘oar’

MG-dialectal κουππί, κουπ-πί, κουπʰι, κουπ-πί, κουπφί, κουπ-πφί

[kuˈp:i], [kuˈpʰi], [kuˈpʰi], [kuˈp:i] (?)

One could of course decide to reject all data not corroborated by modern phonetic research. But this step would mean throwing out about 80% of the collected material, which in most cases is unique or rare and hard to find, and comes from dialects nowadays close to extinction. The dialectal archive of the ILNE is by far the richest in Greece, containing material from the beginning of the 19th century and collected throughout the 19th and 20th centuries thanks to fieldwork by ILNE researchers as well as private individuals (mostly schoolteachers) offering their research to the ILNE archive. Furthermore, even if this material is unreliable phonetically, it still provides invaluable morphological and semantic information which cannot be ignored; it must somehow be included in the Historical Dictionary.

Due to the inclusion of such unreliable data, it has been decided that the ILNE cannot do completely away with its conventional orthographic system. Thus all forms are given both in conventional orthography, with a few extra symbols, and in phonetic

22 The double <σσ> in the Standard MG orthography of the word is simply historical and does not represent geminate pronunciation; in dialectal transcriptions, if the transcriber does not opt for the use of the intervening dash between the two <σ>, it is impossible to tell whether his informant pronounced the word with a geminate [s:] or not. In the next lexical item, the dash is not necessary as the historical orthography never involves double <φφ> (AG did not have geminate fricatives).
transcription. The examples and citations are only given conventionally, without attempting an IPA transcription. An additional reason for maintaining conventional orthography is to avoid making the ILNE utterly incomprehensible to readers. As a national Historical Dictionary, it is not addressed only to the linguistic community, which is familiar with the IPA, but also to students, philologists, historians, specialists in folklore studies and, to a certain extent, to the general public. Of course, the ILNE is incomprehensible to anyone without any philological training at all and without some background knowledge of Greek linguistic history. But one would not in any case expect such users to even be aware of the existence of the ILNE.

In order to transcribe the bewildering variety of amateur phonetic symbols occurring in the older sources as consistently as possible, one of the preparatory tasks of the ILNE was to compile correspondence tables for the symbolism used by the major dialectological studies for each dialect, and to set down strict rules for the use of the custom-made extra symbols complementing conventional orthography. In Table 2, part of the comparative table for the Pontic dialect is given:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>¨α</td>
<td>ia</td>
<td>á</td>
<td>α</td>
<td>α</td>
<td>α</td>
<td>α</td>
<td>æ</td>
</tr>
<tr>
<td>υ, ø, ε</td>
<td>ü</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>ι</td>
<td>w</td>
<td></td>
</tr>
<tr>
<td>χ</td>
<td>χ</td>
<td>χ</td>
<td>χ</td>
<td>χ</td>
<td>χ</td>
<td>χ</td>
<td>j</td>
</tr>
<tr>
<td>ι</td>
<td>ι</td>
<td>ι</td>
<td>ι</td>
<td>ι</td>
<td>ι</td>
<td>p</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Correspondence table for the various phonetic representations of the Pontic dialect

In the following example (12), the old and the new transcription practice of the ILNE are compared:

---

23 The ILNE endorses the view of Abercrombie (1954: 234), according to whom dialectal lexicography should be reader-friendly and include conventional orthographic along with phonetic transcriptions.

24 On the educational role of historical lexicography in society and the concomitant necessity of accessibility to a wide range of disciplines see Reichmann (2012: ch. A.4).

25 The rules are set out in ILNE-MR (2012) and the correspondence tables are available in Manolessou, Beis & Bassea-Bezantakou (2012).
Example from Vol. 4b (1980), lemma γέλιο [ˈʝeʎo] ‘laughter’: only conventional orthography with a few special symbols such as the intervening dash, the hyphen, the hacek and the strikethrough. Each form is followed by abbreviations of place-names where it is attested.

Example from Vol. 6 (forthcoming), lemma Δεκέμβριος [deˈcemvrios] ‘December’: conventional orthography with special phonetic symbols, followed by IPA transcription and abbreviated place-names.

3.3 Databases and digitization

The archive of the ILNE consists of roughly 4 million card-slips, which have resulted from the manual excerpting of manuscript and print sources. Manuscript sources are transcriptions of oral interviews with dialect speakers, mostly made during
fieldwork by the researchers of the ILNE, and dating from the middle of the 19th century (the oldest manuscript dates from 1854). Print sources can be divided in primary material (collections of folktales, songs, etc. from various areas, as well as standard MG texts such as novels, poetry, newspapers etc.) and secondary material (linguistic treatises concerning a dialect or a phenomenon or a lexical item, dictionaries, etc.).

Currently, the manuscript archive is undergoing digitisation which will ensure preservation and ease of access. Major print sources are also being digitised: for example, the relevant section for the letter which is being currently worked on at the ILNE (Delta) has been scanned and digitised from more than two hundred dialect dictionaries and glossaries, so that each lexicographer has very easy access to most of the material. A searchable database containing the lemmata and variant forms for all the letters of the alphabet has also been elaborated, as well as smaller databases for the place-names (exact geographic location, abbreviations), the bibliography (manuscript and printed sources) and for derivational suffixes.

4. Conclusion

The ILNE is a major historical lexicographical enterprise aiming to conform to international lexicographical standards. The delays in its completion are up to a point justifiable due to the huge amount of data that need to be processed. The modernization efforts of the last 5 years have brought to the forefront a number of important methodological issues (data coverage, inclusion of dialectal data, transcription, dating, digitization etc.). The argumentation developed in the present paper and the solutions proposed were meant as a contribution in the metalexicographical literature on historical and dialectal lexicography.

26 The sources of the ILNE and its excerption programme are discussed in detail in Konstantinidou & Tzamali (2012).
References

der der Österreichischen Akademie der Wissenschaften.
BASSEA-BEZANTAKOU, Christina (2010) “The Research Centre for Modern Greek Dialects -
Historical Dictionary”, in A. Ralli et al. (eds.), Proceedings of the 4th International
University of Patras, 10-15.
BASSEA-BEZANTAKOU, Christina (2012) «Λεξικογραφική προσέγγιση της σημασίας» [A
lexicographical approach to meaning and sense], Lexicographicon Deltion, 26, 223-278.
Crusca to the web”, International Journal of Lexicography, 17, 357-384.
BROWNING, Robert (1982) “Greek diglossia yesterday and today”, International Journal of the
Sociology of Language, 35, 49-68.
Routledge.
Kazazis (ed.), The Lexicography of Ancient, Medieval and Modern Greek Literature,
Thessaloniki: Center for the Greek Language, 207-220.
CHATZIDAKIS, Georgios N. (1905-1907) Μεσαιωνικά και Νέα Ελληνικά [Medieval and Modern
Greek], Athens: Sakellarios.
zur Bewertung der sogenannten Archaismen”, in H. H. Munske et al. (eds.), Deutscher
Wortschatz. Lexikologische Studien. Ludwig Erich Schmidt zum 80. Geburtstag von seinen
Marburger Schülern, Berlin/NY: De Gruyter, 525-552.
CONSIDINE, John (1998) “Why do large historical dictionaries give so much pleasure to their
owners and users?”, in Thierry Fontenelle et al. (eds.), EURALEX ‘98: Papers Submitted
to the Eighth EURALEX International Congress on Lexicography, Liège: University of
Liège, 579-587.
<http://dwb.bbaw.de/tagung09/pdf/Durkin.pdf>


<http://www.philology.uoc.gr/conferences/6thICGL/ebook/a/manolessou&giakoumaki &karantzi.pdf>


MANOLESSOU, Io & Nikolaos PANTELIDIS (2011) “Το αρχαίο ελληνικό /e/: στην Ποντιακή” [Anc. Greek /e/: in Pontic], Νεοελληνική Διαλεκτολογία, 6, 245-271.


MANOLESSOU, Io (2012b) “Το Ιστορικό Λεξικό της Νέας Ελληνικής ως ιστορικό λεξικό” [The Historical Dictionary of Modern Greek as a historical dictionary], Lexicographicon Deltion, 26, 9-76.


OECOMNIDES, Demosthenes (1908) Lautlehre des Pontischen, Leipzig: Deichert.


©Universitat de Barcelona


SYMEONIDIS, Charalambos & Dimitris TOMBAIDIS (1999) Η σημερινή διάλεκτος της Ουκρανίας (περιοχής Μαριούπολης) [The modern dialect of Ukraine (Mariupol area)], Athens: Epitropi Pontiakon Meleton.


VALAVANIS, Ioannis (1892) Ζώντα μνημεία τῆς ἀνά τον Πόντον ἱδιωτικῆς [Living monuments of the idiom of the Pontus], Athens: Perris.

Dictionaries


ILNE: (1933-) Ἰστορικὸν Λεξικὸν τῆς Νέας Ἑλληνικῆς, τῆς τε κοινῆς ὁμιλουμένης και τῶν ἰδιωμάτων, 1-5 (α-δαχτυλωτός) [Historical Dictionary of Modern Greek, both of the common spoken language and the dialects], Athens: Academy of Athens.


Kriaras: Emmanouil KRIARAS (ed.) (1968-) Λεξικό της Μεσαιωνικής Ελληνικής Δημώδους Γραμματείας 1100-1669, 1-18 (α-ραβέντι) [Dictionary of Vernacular Medieval Greek Literature 1100-1669], Thessaloniki: Center for the Greek Language.


