

APPENDIX A. HISTORY & DIATOPY OF GREEK

Greek is an Indo-European language with a long history of textual attestation. All the same, the avoidance of the vernacular for most of the past two millennia by the learned classes has made the attestation of the language people actually spoke patchy for significant stretches of that time. In the following section, I give a brief history of Greek, from antiquity until modern times. Then, I review the various dialects of Modern Greek (§A.2), and the Balkan Sprachbund as a factor influencing the development of Modern Greek (§A.3). Finally, I discuss the lacunae in the coverage of my dialect corpus (§A.4).

A.1. A brief history of Greek

Ancient Greek¹ extends from the first attested texts to around 300 BC; the Greeks themselves arrived in the area some time in the second millennium BC, and there is a significant non-Greek substratum in the language which has been argued (Palmer 1980) to have been Anatolian Indo-European. The language was characterised by a great dialectal heterogeneity. There are four dialect groups recognised for Ancient Greek: *Arcado-Cypriot* (including the archaic dialects spoken in Arcadia and Cyprus), *Aeolic* (including Lesbian, Boeotian, and Thessalian; Aeolic is frequently conflated with Arcado-Cypriot, and termed *Central Greek* or *Achaean*), *West Greek* (consisting of *Doric* and *North-Western Greek*), and *Attic-Ionic* (consisting of *Ionic* proper, and *Attic*, the dialect of Athens, closely related to Ionic).

The earliest Greek texts are the Mycenaean texts, written in Linear B script around 1400 BC. The language of these texts is closest to Arcado-Cypriot, and the consensus is that Arcadian was a relic dialect, encircled by Doric-speaking settlers in the Peloponnese. The Linear B texts are almost exclusively statements of accounts, and do not have a rich textual content; nevertheless, although Mycenaean Greek is not Proto-Greek (there is ample internal evidence that the dialects had already diverged by that time), some of the reconstructed features of Proto-Greek can be seen at work in the texts.²

¹For a good summary of the history of Ancient Greek, see Palmer (1980), on which this discussion is largely based.

²Notably the transition from Proto-Indo-European /j/ to Ancient Greek /h/, and of Proto-Indo-European /k^wi/ to Attic-Ionic /ti/. For example, the Classical subordinator *hóti* turns up in Mycenaean as *jo-k^wi*.

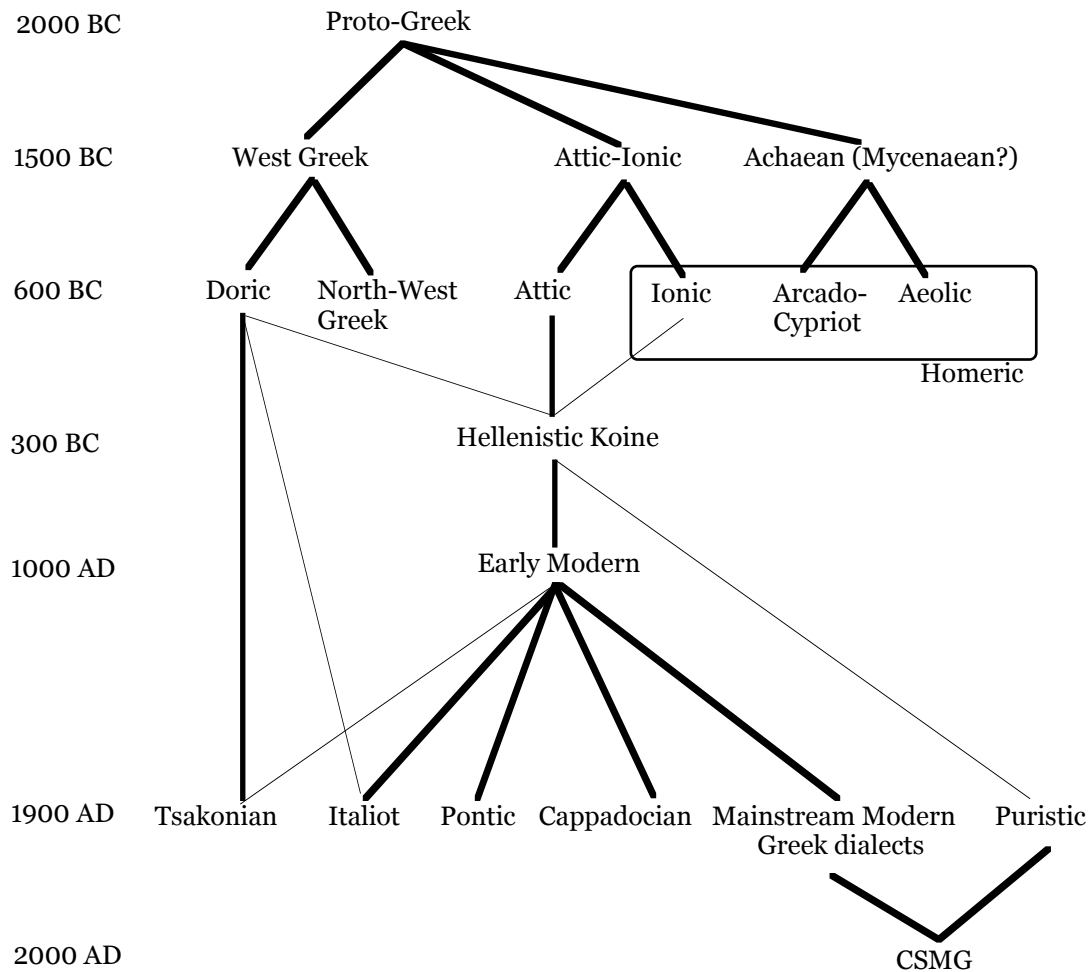


Figure 28. Filiation of historical variants of Greek. Thin lines indicate a minor contribution to koineisation.

Homeric Greek is the language of the Greek Epics, the *Iliad* and the *Odyssey*; these date from late *viii* BC and early *vii* BC, respectively. Homeric Greek is a literary koine composed of various dialects. The language used is basically Ionic, but it is built on a very long history of epic poetry, probably dating back to Mycenaean times. As a result of the oral transmission of the poems, there is a significant Aeolic component to Homeric language, and a smaller Arcado-Cypriot component (which may simply be proto-Achaean, and thus old Aeolic.) Homeric Greek is thus highly heterogeneous, and unlikely to reflect a spoken vernacular; as Palmer (1980:87–88) puts it,

one suggestion may be ruled out at once: the notion that such a mixed dialect represents the spoken language of any historical Greek community... It is hard to imagine a living dialect which possessed three different genitive forms, such as -οιο, -οο, ου or so many different forms of the personal pronouns as ἄμμεζ/ἡμεῖζ [ámmes/hemeîs: ‘we’], ὕμμεζ/ὕμεῖζ [úmmes/humeîs: you.PL], etc.

The Homeric text also displays hypercorrections, misreadings of obsolescent dialect forms, and artificial forms. Nevertheless, its antiquity and length make it a very important witness of Greek.

Homeric Greek is of interest with respect to *pu* because it shows two other relativisers being reanalysed into complementisers (and somewhat later (Rijksbaron 1976) temporal and causal connectives), in a manner parallel to *hórou*; these are *ho:s* (originally ‘the way in which’), and *hóti*, which (as *oti*) has survived into CSMG as a high-register complementiser, and whose temporal and causal function is still active in dialect (originally, ‘whatever’) (§5.3).

Subsequent literature was produced in the local dialects; as different genres originated in different parts of the Greek world, different dialects became associated with them, although the influence of Homeric Greek is always appreciable. Lyric poetry, for example, most famously associated with Sappho, was written in Sappho’s native Aeolic; choral poetry, developed by the Dorians, was written in Doric.³ However, the best preserved and best known works of Ancient Greek—those to which the label ‘Classical Greek’ properly applies—are written in the dialect of Athens, Attic. The language of the orators and comedies is closest to Attic proper; the language of the tragedies is Attic coloured by Homeric, while the language of Classical prose eliminates some Attic phonological features in favour of their General Ionic equivalents. This is the variant of Ancient Greek which has drawn the most scholarly attention.

A major linguistic break in the history of Greek was the formation of Hellenistic Koine⁴ Greek in late *iv* BC⁵—triggerred by Athens’ political dominance, and accelerated by the new political realities after the campaigns of Philip II of Macedon and his son Alexander the Great. Koine was mainly based on Attic, but had significant admixture from Ionic, and some elements of Doric. Koine Greek had pretty much displaced all dialects of Ancient Greek by *i* BC (see Bubenik (1993) for a comprehensive account of the process, drawn from epigraphic evidence.) Doric experienced a revival in inscriptions in Sparta up to *ii* AD (where it is called Neo-Laconian), with varying degrees of success, but researchers have concluded the revival was largely artificial (Panayiotou 1993). Koine is the form of Greek ancestral to all dialects of Modern Greek, with the exception of Tsakonian; it is the language of the Septuagint (the Greek translation of the Jewish Scriptures), the Christian Scriptures and early Christian works, a few literary works (notably Epictetus), and the Egyptian papyri.

From *i* AD onward, Greek writers rebelled against the diffuseness they perceived in literary Koine, and by *ii* AD a thorough-going program of Atticism—emulating the Attic dialect of classical literature—was established. This was the

³Although by the time of the Athenian tragedians, the Doric of the choruses had become quite superficial; Palmer (1980:132) speaks of the “essentially conventional and decorative nature of the literary Doricisms.”

⁴To forestall confusion: *koine* refers to any linguistic variant formed out of dialect merger (koineisation); it is named after the Hellenistic Greek instance (‘common’), which is here capitalised as *Koine*.

⁵The authoritative English-language account of the history of Greek from Koine onwards is Browning (1983); see also Tonnet (1995 [1993]) and Triandafyllidis (1981 [1938]), and more recently Horrocks (1997).

beginning of the phenomenon in Greek broadly described as ‘diglossia’.⁶ High literature was written in Atticist Greek (frequently hypercorrect in its pursuit of the non-colloquial), and when Christianity became a state religion in *iv* AD, the official church came to use Atticist Greek too. Technical prose and official writing was done in Literary (or High) Koine (also called *Fachprosa*—Specialist Prose). Low Koine was an approximation of the spoken language (although everyone who wrote did try, somehow or other, to ‘correct’ their language); it is found in the papyri (which peter out by *viii* AD), and in some of the less pretentious ecclesiastical works.⁷

Historians of Greek (for instance Jannaris 1897) typically split this period as follows:

<i>Hellenistic Greek</i>	300 BC–150 BC
<i>Roman Greek</i>	150 BC–300 AD
<i>Transitional Greek</i>	300 AD–600 AD
<i>Byzantine or Early Mediaeval Greek</i>	600 AD–1100 AD

The onset of Modern Greek is dated at 1100, when literary production in the vernacular commences. But all four periods form part of the one continuum; no major linguistic catastrophes occurred at 150 BC or 300 AD, and the language continued developing until, by *vii* AD, it had acquired most of the features of Modern Greek. It is also clear (see discussion in Browning 1983) that the decisive break between Hellenistic Koine and Modern Greek occurred not at 1100 but in the Greek Dark Ages (650–850)—a period marked by Arab and Slav invasions and religious internecine strife (the Iconoclasm controversy). Indeed, the Protobulgarian inscriptions (*viii* AD) (Beševliev 1963) are already in essence Modern Greek. Since however extensive texts only appear after 1100 AD, I follow this date as the starting point for Modern Greek.⁸

The language spoken between 300 BC and 1100 AD is usually called *Hellenistic*, *Late*, *Post-Classical*, or *Koine*, although ‘Hellenistic’ properly applies only to the pre-Roman period, ‘Late’ and ‘Post-Classical’ do not clearly distinguish this period from Modern Greek, and Hellenistic Koine is neither the first nor the last Greek koineised language variant. For that reason, I use the term *Middle Greek* in this work, by analogy to Middle English, Middle Persian, etc.

⁶While ‘diglossia’ is a convenient cover term to use for the radical social or intellectual stratification of language characteristic of Greek, frequently there are not two but three distinct identifiable variants (as was the case in Roman Greek and Middle Byzantium), or the variants formed a continuum (Puristic/Standard Demotic/Dialect, Puristic/‘Mixed’/Demotic). So I use the term ‘diglossia’ with a degree of looseness, to refer to an *n*-way social stratification of language.

⁷Of particular importance are the *Apophthegmata Patrum* (Sayings of the Desert Fathers) (~480), Malalas’ *Chronographia* (525–550), and the writings of Leontius of Neapolis (~650).

⁸“We are very much worse placed to trace the evolution of the Greek language than during the period of the Roman empire. It is for this reason, amongst others, that the period of political breakdown and demographic change is not considered on its own, but as part of a longer period, during the second half of which we are better off for evidence. But it must be remembered that many of the changes which are first attested in the second half of the period probably occurred during its turbulent first half.” (Browning 1983:55)

The term has already been used (mainly in German, as *Mittlegriechisch*), to refer to Greek spoken in the Middle Ages (500–1500 AD); I use the designation *Mediaeval Greek*, when it is necessary to lump together Late Middle (*Byzantine* in Jannaris' taxonomy—600–1000) and Early Modern Greek (1000–1500).

The linguistic features of Modern Greek not realised by the onset of the Dark Ages were realised during the ensuing upheaval. In this time, the loss in educational opportunities meant that little Atticist prose was written. This did not mean that prose started being written in the vernacular; it was rather written in High Koine, which by this time differed enough from the spoken language that it had to be learnt at school (Browning 1978).⁹

In x AD, the Byzantine empire experienced a renaissance of learning under the Macedonian and Comnenan dynasties. Atticism returned reinforced, as written by authors such as Anna Comnena and Michael Psellus.¹⁰ The period studiously avoids the vernacular; its scholars can hardly be begrudged for this, since they were busy reconstructing Classical learning, and the first surviving manuscripts of most Classical Greek works date from this period. At the end of this period, the first experiments in vernacular literature were made;¹¹ the language used in these poems is macaronic, but identifiably Modern Greek.

In 1204, the Byzantine Empire fell to the soldiers of the Fourth Crusade, and French and Italian started exerting an appreciable effect on the spoken language. The dissolution of the cultural institutions of Byzantium and the model of Western literature allowed vernacular writing in Greek to proceed.¹² These form the bulk of our EMG texts. From xv AD up to xvii AD, there is a steady stream of vernacular text, particularly in Western-held territories—Rhodes, Cyprus, Crete, Corfu. The literature written in these areas is in the local dialect. Prose from the now Turkish-occupied mainland is less plentiful; all through this time, the learned wrote in Atticist Greek, and the advocates of the vernacular were few and far between.

The history of Greek since the establishment of the Modern Greek state in 1833 is tied up with the phenomena of koineisation and diglossia; CSMG has a

⁹The works of the period which admit the most elements of the spoken language include Constantine VII Porphyrogenitus' works (~950), as well as the Proto-Bulgarian inscriptions (viii–ix AD) (Beševliev 1963; 1971), which are arguably our first Modern Greek texts.

¹⁰Indeed, many Lives of Saints were translated at this time from Low Koine into a more acceptable idiom by Symeon Metaphrastes ('The Translator'), and the originals discarded.

¹¹Michael Glykas' *Prison Verses* (1158/1159), the *Ptochoprodromos* cycle (1170s?), possibly a vernacular version of the *Epic of Digenes Akrites*, and the scattered quatrains written by the Turco-Persian poet Mevlânâ Rumi and his son Sultan Veled (xiii AD). By vernacular, I mean displaying substantial modernisms in grammar unparalleled in the literary Greek hitherto used.

¹²Major vernacular texts written during this time include the vernacular romances (Beaton 1996), the *Assizes of Cyprus* (~1250)—a legal code heavily influenced by Old French, the *Chronicle of Morea* (1370s/1388), and the Cypriot-dialect *Chronicle of Makhairas* (~1432). There is also a wide variety of mainly moral or didactic poetry whose date is uncertain, but which can be traced to xiv–xv AD.

significant Puristic¹³ component, and is a dialect koine based mainly on Peloponnesian.

A.2. A dialect survey of Modern Greek

At the turn of the century,¹⁴ a great variety of Greek dialects was spoken in Italy, the Balkans, and Asia Minor (see Map 5.) Most of these dialects are now either dead or moribund: population exchanges, increased mobility, compulsory military service, massive urbanisation, universal education and the mass media have combined to restrict dialects to the rural elderly, although some dialect colouring persists in regional variants of CSMG. The only current exceptions to this trend are Cypriot, and to a much lesser extent Cretan; yet even in these areas, CSMG is in widespread use.¹⁵ Contossopoulos (1985 [1981]:87), a specialist on Cretan, is candid in discussing the relevance of dialect to the wider Greek community:

Especially concerning Crete, I do not believe dialectological public lectures are possible any more. The topic has been done to death and—why fool ourselves—perhaps by now, in speaking of dialects, we offer any audience not consisting only of specialist linguists something of a linguistic illusion, so to speak. We are speaking of bygone matters. The more elderly have a clearer image of the phenomena the dialectologist speaker will go through. But the younger will ask themselves: ‘Were such things really ever spoken in our region?’ The villagers amongst them will wonder: ‘Is that *nonsense* what scholars spend their time on?’ Some younger people with ‘literary interests’ or something of the sort will perhaps make a note of some grammatical phenomenon or some peculiar local word (as a linguistic curio) to present in their written, or even their oral discourse as a ‘special effect’, within the framework of a more general ‘return to roots’, a cultural ‘retro’.

Contossopoulos concludes his paper by declaring:

the time for the description of dialects and dialect phenomena is over. The dialects are extinguished or are being extinguished, and dialect phenomena are now difficult to follow. The time has come to research *what* has remained of the dialects, *why*, and *how*. (Contossopoulos 1985 [1981]:92)

With one or two exceptions, all Modern dialects are descended from Middle Greek Koine. The date at which the dialects diverged from Koine is not certain. Dialect-coloured texts exist from *xiii* AD, but dialect does not feature regularly in Greek texts until *xv* AD (Cypriot and Cretan). For the more archaic dialects (Cypriot, Cretan, Pontic, Cappadocian), Browning (1983:130–131) considers the Saracen invasions of Byzantium (*vi–ix* AD) a likely starting-point for divergence.

¹³**Puristic** (*katharevousa*) is a neo-classical linguistic variant of Greek, based on older Greek (mostly Hellenistic); it was the official language of the Greek state until the 1970s, and was intended to spearhead the revival of Ancient Greek. The vernacular, in opposition to Puristic, was called **Demotic**.

¹⁴Notwithstanding the fact that this work is itself written at the turn of a century, I use ‘turn of the century’ to refer to the period around 1900.

¹⁵Dialect death is slightly retarded in the Greek diaspora, where the pressure to conform to a standardised Greek is attenuated—but in a context where the eventual death of the language itself is assured.

There is evidence for characteristic dialectal phenomena dating back to *vi* AD (Babinotis 1976; Delopoulos 1983 [1978]; Dieterich 1970 [1898]; Macharadse 1980).



Map 5. Greek-speaking regions, 1900

The exceptions to the descent of Modern dialects from Middle Greek Koine are Tsakonian, and to a lesser extent Italiot. Tsakonian is closer to the ancient Doric dialect than to Hellenistic Koine in its phonology, morphology, and vocabulary; its syntax, however (especially in the latter part of *xx* AD) is much closer overall to the standard language. There is enduring controversy on the extent to which Italiot preserves features of Doric; the likeliest interpretation is that it is descended from a Doric-coloured regional variant of Koine, although it did not become isolated from the remainder of the Greek-speaking world until *xi* AD, when Southern Italy was lost by the Byzantine Empire to the Normans.

A distinction is made in Greek dialectology between three classes of dialects: those not mutually intelligible with CSMG, those mutually intelligible, and an intermediate class. Greek dialectologists prefer to restrict the term *dialect* to the first and third class, terming variants mutually intelligible with SMG *idioms*. Following practice in English, the term **dialect** is retained throughout here to refer to regional variants of a language; those forms not mutually intelligible with Greek are termed **outlier dialects**.¹⁶

In the following, I briefly sketch the most important Greek dialects. In examining the distribution of *pu* in Modern Greek, I investigate all dialects on which I have data; this sketch is intended to help situate that data within a diatopic context.

A.2.1. Dialects not mutually intelligible with CSMG

There are four variants of Greek not mutually intelligible with CSMG, and they are termed dialects rather than distinct languages for cultural rather than lin-

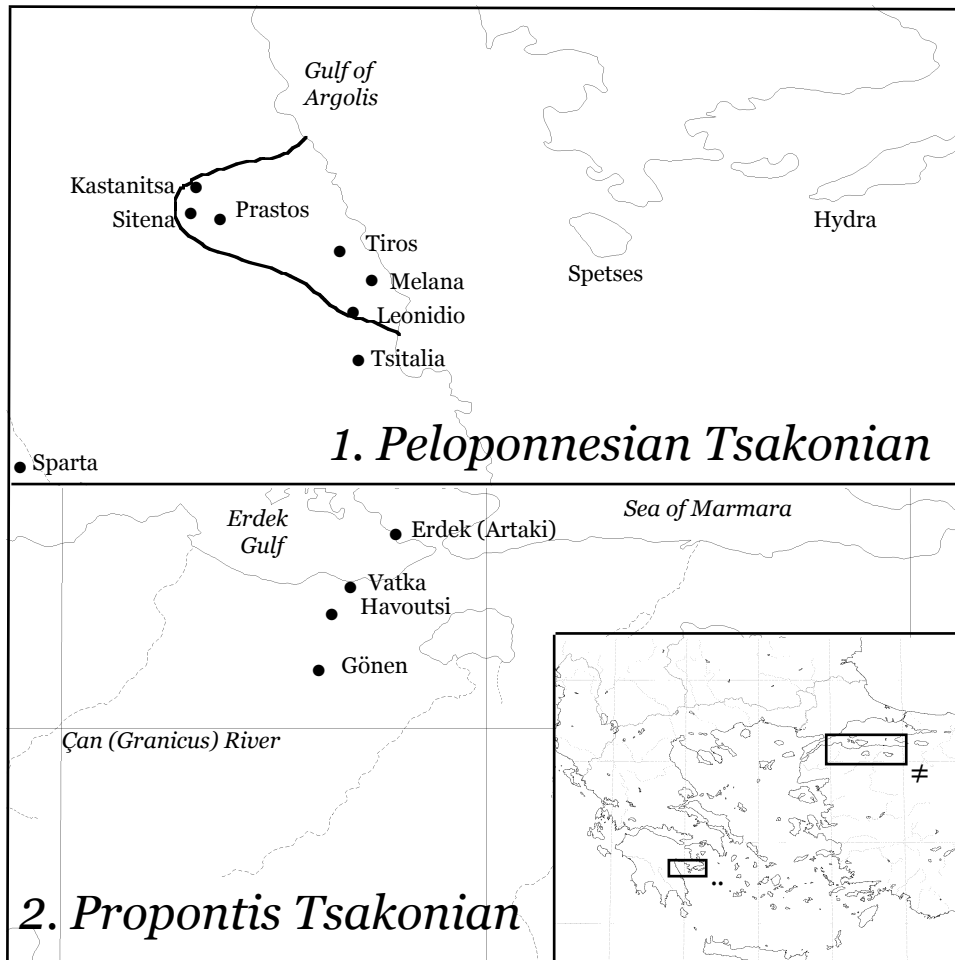
¹⁶This designation was used by Triandafyllidis (παράμερα ιδιώματα) (1981 [1938]:69), who applied it however to refer only to the geographically peripheral dialects—i.e. Pontic, Cappadocian, and Italiot, but not Tsakonian.

guistic reasons. **Tsakonian** is very much *sui generis*. The other three were or are spoken outside the contiguous Greek-speaking sphere, and display extensive language contact alongside their archaisms: **Calabrian** and **Apulian Italiot**, in Southern Italy, and **Pontic** and **Cappadocian** in northern and central Turkey.

Tsakonian

Tsakonian (Costakis 1951; Costakis 1986; Haralambopoulos 1980; Pernot 1934) is the most archaic Modern Greek dialect; non-Greek linguists usually refer to it as a distinct language. Tsakonian was identified very early as not mutually intelligible with SMG; Gerlach wrote to Martin Crusius in 1578:

and almost all the differences [between Greek dialects] are in pronunciation and a few words. And all of them, from whichever region, understand each other, except for the Ionians who inhabit the Peloponnese... These are called in the vernacular Tsakonians (cited in Triandafyllidis 1981 [1938]:428).



Map 6. Regions where Tsakonian is/was spoken

Tsakonian has three subdialects (see Map 6). The dialect as spoken in the Peloponnese is subdivided into **Northern**, spoken in the villages of Kastanitsa and Sitena, and **Southern**, spoken by the remaining villages in Tsakonia in the

Eastern Peloponnese; there is also **Propontis Tsakonian**, spoken by Tsakonians who migrated to the villages Vatika and Havoutsis on the Sea of Marmara (= Propontis) some time between *xii* and *xviii* AD,¹⁷ and heavily influenced by the more mainstream Thracian dialect of Greek. All Tsakonian speakers are now bilingual with CSMG; it is doubtful that the language is currently spoken by any more than 300 people. Tsakonian has a radically different phonetic and phonological system to CSMG, and its morphology has evolved quite independently of Modern Greek.

Tsakonian appears to be descended from the ancient Doric dialect, whereas all other Greek dialects are descended from Hellenistic Koine, which was largely Attic-Ionic.¹⁸ Fixing a date on the divergence of Tsakonian from Standard Greek is problematic. If we postulate that Tsakonian is directly descended from Doric, without any Koine admixture, the divergence should be dated at the break-off of Doric from Proto-Greek—some time in the second millennium BC. But this is unreasonable: at every turn, Tsakonian displays evidence of close contact with the Greek mainstream. Phonologically, the split must be dated at around 800; the grammatical data is ambivalent, as SMG innovations are present in Tsakonian, but several innovations made before *vi* AD are absent in the language.¹⁹

The historical endpoint for the divergence of Tsakonian from Standard Greek seems to be the Slavic invasions of the Peloponnese in *viii* AD. Tsakonians were frequently conscripted in the Byzantine army, and were known of in Constantinople.²⁰

¹⁷Vatika had a population of 323 in 1913, and Havoutsis 171 in 1922 (Costakis 1979:30); since 1922, this population was resettled in Greece, and the dialect has now died out.

¹⁸There was some doubt about the Doric parentage earlier this century; the filiation of Tsakonian is no longer disputed. Nonetheless, the influence of Standard Greek on Tsakonian is pervasive, and the possibility of earlier non-Hellenic influence on the morphology has not been conclusively refuted—although there is no substantial evidence for it in the lexicon.

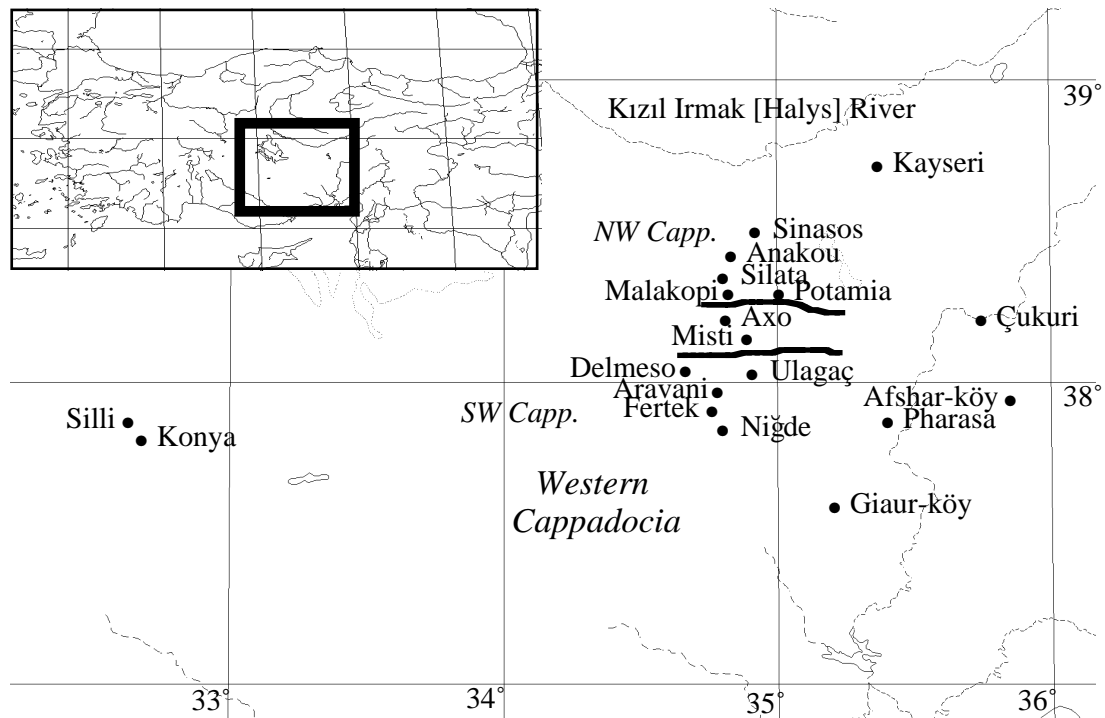
¹⁹For further discussion, see Nicholas & Sidwell (1997). The discussion in that paper is couched in terms of lexicostatistics, using the Swadesh-100 list. The Swadesh-100 and 200 lists are lists of core content-word vocabulary (Gudschinsky 1956), through which linguistic divergence has been attempted to be quantified in lexicostatistics. Lexicostatistics remains controversial, particularly in its use to also date divergence (glottochronology); however, it is useful as at least an indicative quantification.

In terms of the Swadesh-100 list, Southern Tsakonian has 70% of basic vocabulary cognate with CSMG, but at least 8 of the 70 cognates are recent, phonologically unassimilated loans from CSMG. If the glottochronological constant of 86% retention per millennium of the Swadesh-100 list were assumed to be valid (which it generally is not nowadays), this would give a separation date between Tsakonian and SMG of 400±400 AD (90% confidence interval).

²⁰The first explicit allusion we have to Tsakonian not being intelligible by Standard Greek speakers comes in the fifteenth century satire *Mazaris' Sojourn in Hades* (Barry, Share, Smithies & Westerink 1975:65): "I thought I would myself turn into a barbarian, just like the Laconians have become barbarians, and are now called Tsakonians"—although as Pernot (1934:240) points out, the words Mazaris quotes as 'barbaric' are actually Maniot (see below).

Cappadocian

Cappadocian (Dawkins 1916) is a cover-term for dialects spoken in central Asia Minor (Cappadocia) until the 1922–24 population exchanges between Greece and Turkey. They constituted Greek-speaking islands within an increasingly Turkish-speaking area, and were themselves significantly affected by Turkish—to the point where some villages had borrowed vowel harmony and numerals from Turkish, and used substantially Turkish syntax. Dawkins' (1916:198) statement “Turkish has replaced the Greek spirit: the body has remained Greek, but the soul has become Turkish” is widely quoted in the literature as an epigrammatic description of language death.²¹ Cappadocian was retreating not only before Turkish, but also before Standard Greek (in its Constantinopolitan guise), making inroads into the region both through trade and education; Dawkins (1916:3) reports that the local dialect of the town of Sinasos, described in the 1890s, had already yielded to Standard Greek, and Standard Greek was also advancing in Silli.



Map 7. Cappadocia

Dawkins records three groups of Greek dialects in the Turkish hinterland: **Western Cappadocian** (some twenty villages, to the south of Kayseri) (Costakis 1964; Fosteris 1952; Fosteris & Kesisoglou 1960; Kesisoglou 1951; Mavro-

²¹The Turkish–Greek admixture of Cappadocian has attracted attention in language contact studies because of the typological disparity of the two languages—and because an Indo-European language is involved. Morosi’s 1874 statement on Italiot, that “the matter is Greek, but by now the spirit is Italian” (cited in Taibbi & Caracausi 1959:lxxxii), is presumably independent of Dawkins’.

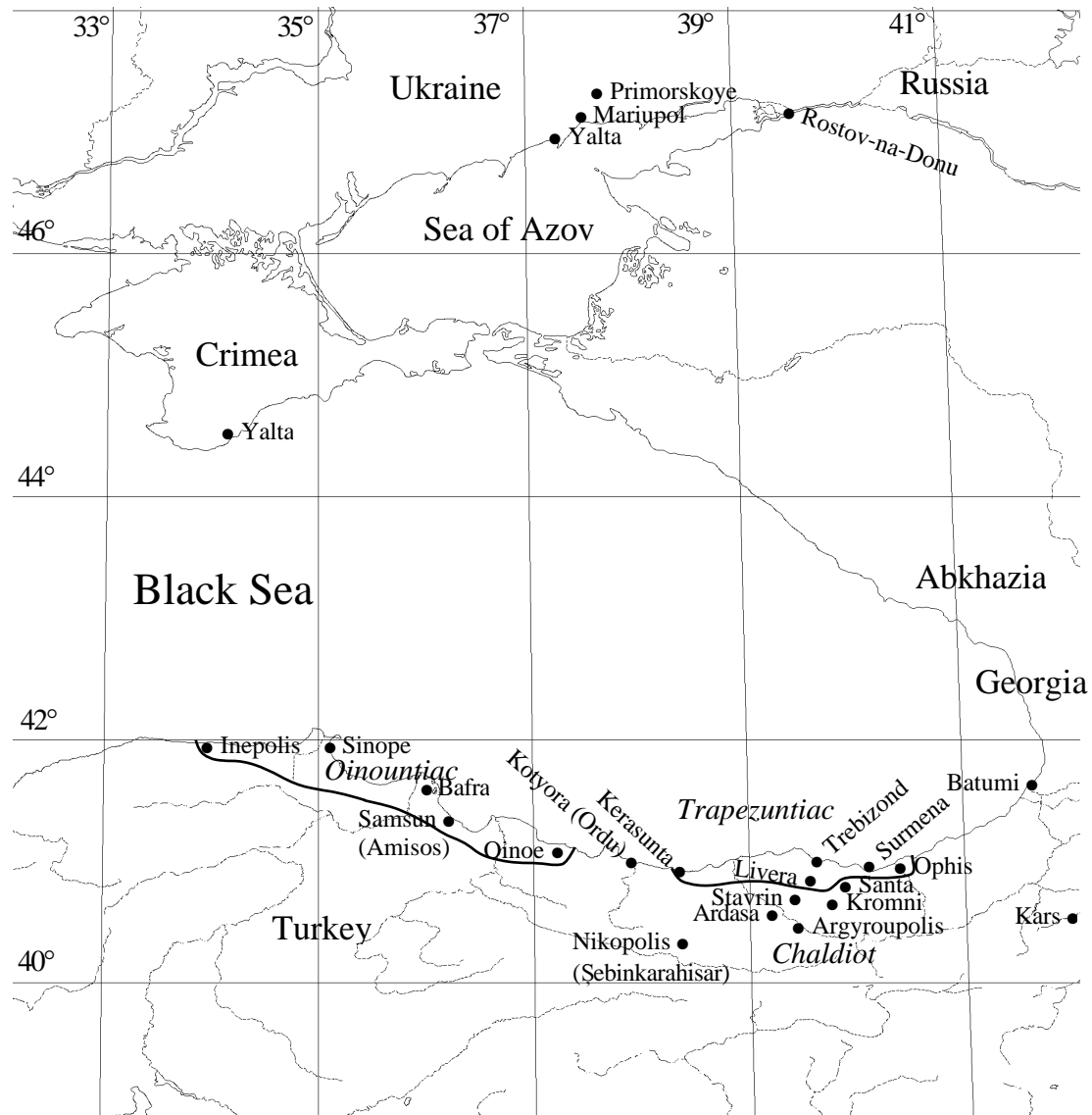
halyvidis & Kesisoglou 1960), the villages around **Pharasa** [Faraş] to the east of Cappadocia (Anastasiadis 1975; Andriotis 1948), and the village of **Silli** [Sille], near Konya, to its west (Costakis 1968). Of these, Pharasa was least subject to Turkish influence, while Cappadocian proper was subject to it the most. According to Contossopoulos (1994), all these dialects have recently died out with the last of the first generation of refugees from Asia Minor.

Cappadocian had been isolated from mainstream Greek (with the exception of Pontic to its north) since *viii* AD, with the Arab conquest of Cilicia to its south (Anastasiadis 1975:154), and even more so after the eleventh century, with the arrival of the Turks (Anastasiadis 1976:3). Dawkins (1916) suspects that Pharasa was originally a Pontic mining colony; its similarities with Pontic are appreciable. In several ways, the dialect of Silli is closer to SMG than the Cappadocian dialects proper; for example, it has a reflex of the future particle *θα* (*se*), absent in Italiot, Western Cappadocian and much of the Pontus, and the recent Passive PERFP suffix *-ka*, again absent in Italiot, Cappadocian, and the Pontus (Dawkins 1940:13). For this reason, Dawkins considers Silliot a remnant of a West–Asia-Minor Early Modern vernacular, related to Mariupolitan (see below). The *xiii* AD Greek poems of Mevlânâ el-Jalal Rumi and Sultan Veled have been argued to bear traces of Cappadocian; the father and son Turco-Persian poets lived in Konya, near Silli.

Pontic

Pontic (Athanasidiadis 1977; Dawkins 1931; Dawkins 1937; Drettas 1997; Ikonomidis 1958 [1940]; Papadopoulos 1955b; Papadopoulos 1961) refers to Greek dialects formerly spoken on the southern shores of the Black Sea ('Pontus'); it is still spoken there, around Ophis [Of], by the Muslim Pontians not subject to the 1922 population exchanges (Mackridge 1987). Pontic was spoken in isolated pockets in the western part of the Pontus, from Inepolis [Inebolu] to Oinoe [Ünye], and in a continuous region from Oinoe to Ophis. It was also spoken in Pontian mining colonies further inland (as seen, Pharasa has been considered an erstwhile mining colony), and in isolated pockets near the Caucasus, such as Batumi and Kars.

Pontic is also spoken on the northern shores of the Black Sea, in southern Russia, Abkhazia, Georgia, and the Ukraine, particularly around Rostov-na-Donu (Semenov 1935); the Pontic speakers are known to have emigrated from Asia Minor in *xviii* and *xix* AD (Dawkins 1937:17–18). Many of the Northern Pontic speakers were resettled during Stalinist rule to Kazakhstan; following the collapse of the Soviet Union, they are resettling in Greece. Their variant of Pontic has not yet been studied extensively. Drettas (1997) gives a count of some 300,000 Pontic speakers in total.



Map 8. Pontus

Pontic is usually also taken as including **Mariupolitan** (Karpozilos 1985; Zhuravliova 1993),²² now spoken around the town of Mariupol (Zhdanov) in the Donets region of the southern Ukraine, and originally spoken in the Crimea, from where the Greeks moved in 1778.²³ Contossopoulos (1994:10) calls it “a mixed idiom with Pontic and northern Greek elements”, and the relationship to Northern Greek has been emphasised by Russian and Ukrainian researchers. However, Dawkins (1937) rejects this classification; he believes it is a continuation of the Greek spoken by ancient colonists to the northern Pontus, and con-

²²Outside the Soviet Union, Mariupol Greek is the name given to the dialect; the autonym of the Ukrainian Greeks is *rumei* ‘Romans’ (the pre-Enlightenment autonym of Greeks generally), and Soviet/CIS linguists refer to the dialect as Crimeo-Rumeic or Tavrurumeic (Zhuravliova 1993:562).

²³In fact, one of the Mariupolitan villages bears the name of Yalta, indicating that its famous namesake on the southern tip of the Crimea was formerly Greek. (A village immediately to the south of the Crimean Yalta bears the unmistakably Greek name of *Livadiya*, ‘Meadows’.)

siders it to resemble the dialect of Silli more than any other. Mariupolitan and Silliot have an important role in helping reconstitute the linguistic picture of pre-Ottoman Asia Minor; the Greek spoken in western Turkey (Bithynia, Ionia) does not help to this end, as it seems to have resulted from latter-day migrations from Greece (Dawkins 1916:5, 37).

Pontic has been divided into three dialect groups by Triandafyllidis (1981 [1938]:290). **Oinountiac**²⁴ was spoken in the western, non-contiguous part of the Pontus, from Inepolis to Oinoe. Dawkins (1931:391) reports that Standard (Constantinopolitan) Greek had prevailed in Sinope [Sinop] by the end of *xix* AD, and was making inroads in all the big coastal towns. Dawkins does not even mention Inepolis as a Pontic-speaking town, and in his grammar of Pontic, Papadopoulos (1955b) was only able to find songs, rather than prose, from the town; furthermore, the Inepolitan texts are the only text sample he does not feel it necessary to gloss into CMSG.²⁵

Trapezuntiac was spoken on the eastern shore of the Pontus, from Kerasunta [Giresun] to Ophis, and is named after the main city of the area, Trebizond (Trapezunta) [Trebzon]. **Chaldiot** was spoken in the Chaldia region, south of the eastern shore, and including Argypolis [Gümüshane; Kanin in Pontic] and its surrounding villages, as well as the southern mining colonies and the coastal town of Kotyora [Ordu in Turkish and Pontic]. Dawkins (1931) considers Chaldiot “may be taken as the type, as being that of the most compact Greek-speaking population.” Pharasiot is closely related to Chaldiot; Dawkins (1937) established that the printed Pontic of Rostov is also Chaldiot.

Papadopoulos (1955b) limits himself to a two-way distinction between Oinountiac and Trapezuntiac–Chaldiot. In the context of the questions asked in this work, this is the most salient dialect split in Pontic, given that Oinountiac tends to pattern more closely with mainstream Greek; I refer to these two variants as **Western** and **Eastern Pontic**.²⁶ As Papadopoulos (1953:90) points out, “the speakers of this idiom constitute almost nine tenths of Pontic Hellenism, and the most numerous of those were the inhabitants of Chaldia”. In other words, Chaldiot was the Pontic norm, and responsible for much of the Pontian colonisation to other regions—including Trapezuntiac and Oinountiac-speaking areas.²⁷ Oinountiac, by contrast, was peripheral, and subject to SMG influence; Ophis and Sourmena are archaic in some key aspects.

²⁴The autonym is *Niotika*, which Triandafyllidis (1981 [1938]:653) preferred in his errata.

²⁵In my corpus, I only have prose texts from Upper Amisos (the old town) and Oinoe.

²⁶In addition, Papadopoulos groups the variants of Trebizond itself and the Chaldiot hinterland together as a single idiom (i.e. he groups Trebizond with Chaldiot, rather than the other coastal towns to its east (Ophis, Sourmena [Sürmene]) or west (Kerasunta, Tripolis [Tirebolu]).)

As Drettas (1997:21) points out, Eastern Pontic is encompassed by the mediaeval Empire of Trebizond.

²⁷For example, according to Papadopoulos, only the old town of Amisos (Upper Amisos/Kadiköy) spoke an Oinountiac dialect; the countryside around it spoke Chaldiot, and the new town had Greek-speakers from throughout Asia Minor.

Pontians have displayed great cultural resilience since their resettlement in Greece, and there exists a Pontic-dialect press and theatre (Lampsidis 1952); but although Pontians still constitute a well-defined ethnic group within Greece, the dialect there is nearly extinct there in communicative use, and its contemporary usage is mostly emblematic, rather than functional.

Pontic has a high level of lexical influence from Turkish;²⁸ it has many morphological features either archaic or innovatory compared to CSMG. Pontic was not as isolated from the Greek mainstream as Cappadocian: the cut-off date for contact with the mainstream is probably closer to the fourteenth century,²⁹ Ponticisms already show up in church deeds dating from *xiii* AD (Lampsidis 1952), and Pontic was already being described as a distinct form of Greek in the 1510s (Dawkins 1937:16).

Dawkins (1916) records the existence of a number of other indigenous Greek-speaking areas in Asia Minor. This includes:

- a number of villages between the Pontus and Cappadocia, in the region of Nikopolis (Şebinkarahisar).³⁰ The language spoken here is reported to have been intermediate between Pontic and Cappadocian; regrettably, no substantial work has been done on the dialect, despite remarks on its significance by Hatzidakis and Dawkins.³¹ We do know that the linguist Papadopoulos (a native of Argypolis) did not find it intelligible (Papadopoulos 1953:85).
- Gölde in Lydia (whose dialect had become extinct by the late nineteenth century.)
- Livisi (Kaya) with its colony Makri (Fethiye), on the southwestern coast—which Dawkins considers an Anatolian dialect like Cappadocian, although Contossopoulos (1994) considers it an idiosyn-

²⁸In addition, there is syntactic influence: Nicholas Contossopoulos, a Greek dialectologist of Turkish-speaking Christian descent, has observed to me that the syntax of Pontic is very similar to that of the first generation of Turkish-speaking Christians who learned Greek.

²⁹See, however, Dawkins (1931:398):

A glance at the map will show us that a great wedge was driven by the Seljuq invasions of the eleventh century between the Greeks of the east and those of the western parts of Asia Minor. The Cappadocian dialects lie at the eastern fringe of this Seljuq empire; Silli is close to its capital [Konya]. The Seljuqs, that is, did nothing to break up the linguistic unity of Cappadocia and Pontos, but a great deal to separate these from anything further west; and first of all from the kind of Greek of which the dialect of Silli was the only example surviving to our own day. The separation of Pontos from the area of the Cappadocian dialects must be put down to later events; to the Ottoman and other Turkish tribes who overran all Asia Minor in the thirteenth century.

One should note, however, that traffic between the Greek Trebizond Empire and Constantinople continued during the Seljuk period by sea; this may well have delayed the linguistic separation of the Pontus from the Greek mainstream, compared to Cappadocia.

³⁰In the town of Şebinkarahisar itself, the Christians spoke Turkish (Dawkins 1931:391).

³¹The Centre for Asia Minor Studies (Κέντρον Μικρασιατικών Σπουδών) in Athens has a small manuscript grammar and text collection (Merlier 1974:225), which I have not sighted. Some four thousand words of the dialect are in my text corpus.

cratic semi-northern dialect (see below), and which also has clear affinities with South-Eastern Greek.

Italiot

Italiot (Karanastasis 1991; Rohlfs 1950; 1962; 1964; 1974) is spoken in southern Calabria and Apulia, in the region of Salento, by at most 20,000 speakers; it is mostly spoken by the elderly, and is fast dying out, at least in Calabria.³² The origins of Italiot³³ have been the subject of much controversy; as a consequence, there is a much greater non-Greek-language literature on these than on any other dialect of Modern Greek. The current consensus is that Italiot is a continuation of a Doric-coloured local variant of Hellenistic Koine, which became cut off from the Greek mainstream with the Norman conquest of the eleventh century.³⁴

Italiot heavily borrows function words as well as content words from the local Italian dialects, Calabrese and Salentino; morphologically, the dialect is very conservative. The most telling evidence for the isolation of Italiot is the absence of grammaticalisations in Early Modern Greek such as *tha*.

The dialect is of considerable diachronic interest—as is also the case with Pontic and Cappadocian—because of its longstanding isolation from Balkan Greek. If a feature of Modern Greek is absent from the most archaic Modern Greek dialects, it can be safely dated as arising after the eleventh century. This is the case with the future particle *tha*, which first appears in its grammaticalised form (< *thelo na* ‘I want to’) in the fifteenth century; it is absent from some Cappadocian and Pontic dialects, and from Italiot (where the grammaticalisation has not progressed beyond an optional volitive future in *te’na*.)

If, on the other hand, the feature is present in Italiot, it was probably present, or at least incipient, in eleventh-century Greek. Furthermore, there is a much smaller chance that it resulted from language contact in the Balkans. (The Slavs are known not to have arrived in the Balkans before the sixth century, although the Albanians are probably indigenous to the region.) It also seems that there has been no contact between the Greek- and Albanian-speaking populations of Southern Italy, the Albanians (Arberësh) settling far from the Greek-speaking regions. So Italiot is a de-Balkanised Greek.

³²Rohlfs (1950) estimates 2,200 speakers in Calabria and 16,000 speakers in Apulia. One may reasonably expect these figures to have dropped drastically in the generation since. Katsoyannou (1994:552) reports that in Calabria the village of Galliciano “is today the only one for which the notion of a linguistic community remains valid, and where I met the youngest Greek-speakers of the region (now between 25 and 30 years old).”

³³*Italiot* is used here as a succinct designation distinct from ‘Italian’. Greek linguists tend to use *κατωϊταλικά* (ελληνικά) (Southern Italian (Greek)), while the autonyms are *Greko* (Calabrian) and *Griko* (Apulian).

³⁴There is limited—and not incontrovertible—linguistic evidence of migration from Greece in the twelfth and/or fifteenth centuries (Karanastasis 1978).



Map 9. Italiot (Salentine, Calabrian)

A.2.2. Dialects barely mutually intelligible with CSMG

Cypriot

Cypriot (Contossopoulos 1969–1970; Loukas 1979 [1865–1898]; Menardos 1969; Newton 1972a) is spoken in Cyprus. It is the healthiest of Modern Greek dialects, for geographical and political reasons; literature and songs are still being written in Cypriot, and Cypriot dialect retains a role in public life (Pavlou 1996). However, both CSMG and Puristic have a significant presence;³⁵ the current linguistic situation in Cyprus is comparable to the diglossic status of *Schwyzerdeutsch* versus High German.³⁶

³⁵Significantly, Cyprus has not experienced the backlash against and removal of Puristic Greek from public life; Cypriots do not feel CSMG is ‘their’ language.

³⁶See Hristodoulou (1973) for a somewhat tendentious summary of the situation.

Cypriot is grouped together with the dialects of the Dodecanese and nearby islands as a South-Eastern dialect; Dawkins (1921) has also claimed the dialect is related to Pontic and Cappadocian. There are phonological differences between various parts of Cyprus (Newton 1972a), now disappearing following the evacuation of the Northern Cyprus Greeks after the 1974 Turkish invasion, and the steady spread of CSMG.

Cypriot has a long history of textual attestation, dating back to the translations of the French Laws (*Assizes*) of the Kingdom of Cyprus (~1250). Cypriot deviates significantly from CSMG in all linguistic domains; Vassiliou (1995), for example, has claimed that Cypriot is a VSO language, as opposed to CSMG, which is SVO.³⁷

Cretan

Cretan (Contossopoulos 1988; Kafkalas 1992; Pangalos 1955) is spoken in Crete, and until recently by the Muslim Cretans who resettled in Turkey this century. There is still some limited literary production in Cretan, and Cretan parochialism ensures the distinctiveness of the Cretan variant of CSMG; but although dialect death has been delayed relative to the rest of the Greek state (Contossopoulos 1985 [1981]), the dialect itself is now fast dying out.³⁸ As Contossopoulos' (1988) atlas vividly illustrates, there are many lexical divergences within Crete, radiating out for the most part from the four main cities of the island, Hania [Canea], Rethymno, Iraklio [Candia/Kastro], and Sitia (Lasithi prefecture).

The differences between Cretan and CSMG are less pronounced than for Cypriot (for which mutual intelligibility is much less), but are appreciable, and extend to syntax as well as lexicon and morphology.³⁹

A.2.3. Dialects mutually intelligible with CSMG

The remaining dialects of Greek are mutually intelligible with CSMG, and differ from it mostly phonologically, and to a lesser extent morphologically; it is for this reason, it seems, that earlier scholars like Crusius and Portius could speak of a common Greek language, without a supraregional variant necessarily being in place.

Dialectal differentiation in Greek occurred because of both historical and geographical factors. For instance, the survival of Tsakonian and Maniot was due to the relative inaccessibility of the regions, and that of Old Athenian due to the isolation of its speakers from the rest of the Greek-speaking world by the surrounding Albanian-speaking population. On the other hand, the distinction be-

³⁷Newton (1972a:110) reports 84% cognates on the Swadesh-200 list, and there are 93% cognates on the Swadesh-100 list.

³⁸See Contossopoulos (1970) for documentation of the dialect death process.

³⁹Lexicostatistically Cretan is much closer to SMG than, say, Tsakonian: the Swadesh-100 list gives 93% common vocabulary, as with Cypriot. The distance between Cretan and CSMG is thus comparable to that between Ukrainian and Byelorussian.

tween Heptanesian and Peloponnesian (such as it is) is largely due to the long-standing Venetian presence in the Heptanesa. Mainstream Modern Greek dialects have been classed into five groups by Newton (1972c), on the basis of the phonological phenomena detailed in Table 24: *Peloponnesian-Heptanesian, Northern, Old Athenian, Cretan, and South-Eastern.*

<i>Phonological Feature</i>	<i>Pel</i>	<i>Nth</i>	<i>O.Ath</i>	<i>Cret</i>	<i>S.E.</i>
Deletion and raising of unstressed vowels	–	+	–	–	–
Stressed [i] does <i>not</i> become glide before vowel	–	–	+	–	–
Reflex of /y/ is [u]	–	–	+	–	–
Velars have (palato)alveolar, not palatal allophones	–	–	+	+	+
/γ/-epenthesis in <i>-evo</i> verbs	–	–	+	+	+
Gemination	–	–	–	–	+
/v, ð, γ/ lost intervocally	–	–	–	–	+
Retention of word-final /n/	–	–	–	–	+
Manner dissimilation (eg. /vy/ > [vg])	–	–	–	–	+

Table 24. Mainstream Modern Greek Dialect Groups (after Newton 1972c)

1. Peloponnesian-Heptanesian (Contossopoulos 1976 [1975]; 1985a). This dialect group includes the Peloponnese and the Heptanesa.⁴⁰ These dialects are considered to have provided the basis for SMG:

Peloponnesian and Ionian dialects have played a crucial role in the evolution of what Greek scholars nowadays term the ‘modern Greek common language’, η νεοελληνική κοινή, so that Peloponnesian dialects in particular differ only marginally from ‘educated Athenian’. Doubtless because of this, Peloponnesian–Ionian has received disappointingly little attention from Greek scholars and apparently none whatever from foreigners. (Newton 1972c:14)

As is clear from Table 24, a significant reason why these dialects became the basis for CSMG is that they are phonologically conservative—without being so conservative as to retain features abandoned by all other mainstream Greek dialects, as was the case with Old Athenian. They are in a sense the phonological ‘lowest common denominator’ of Greek dialects.

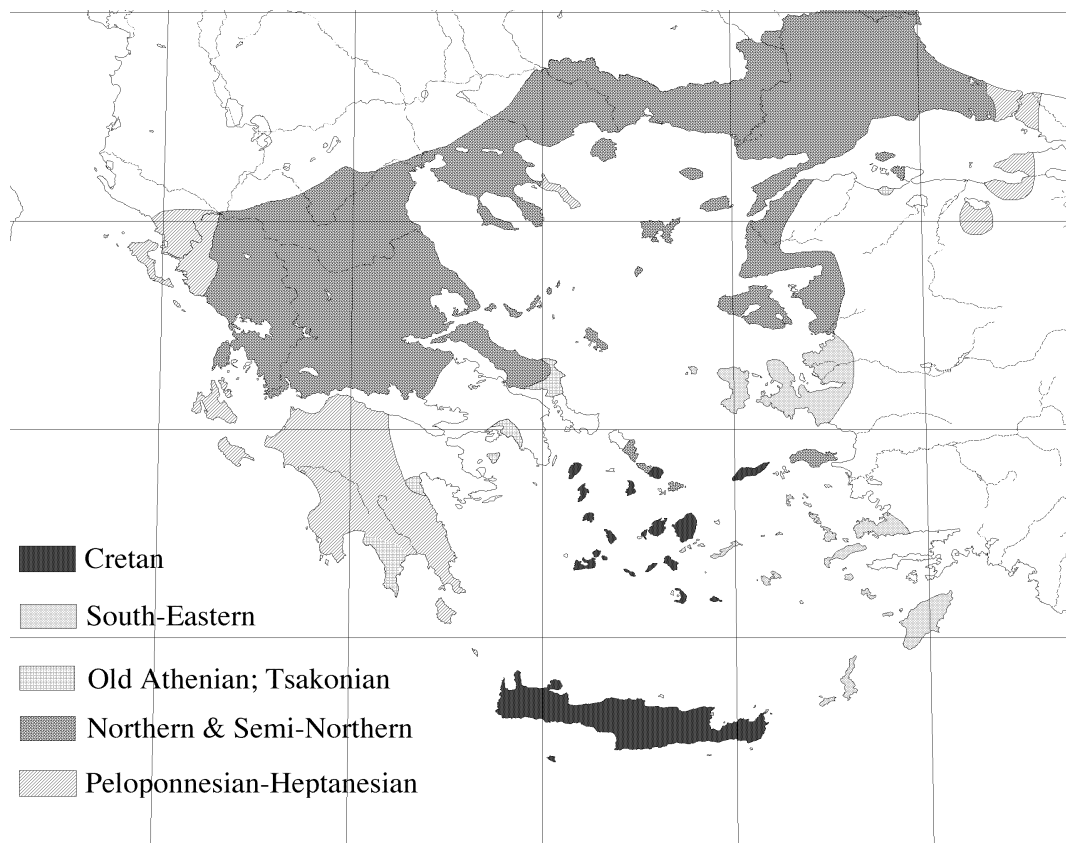
It is because Peloponnesian and Heptanesian have formed the basis of CSMG, and do not differ greatly from it, that these dialects have received the least attention from Greek dialectologists.⁴¹ Dialect studies suffer as a consequence; for

⁴⁰I use the term Heptanesian, rather than Ionian as in Newton (1972c), in order to avoid confusion between the Ionian islands and Ionia (central western coastal Turkey). ‘Heptanesa’ means ‘seven islands’, but as used here it refers to six: Corfu [Kerkira], Paxi, Lefkada, Ithaca, Cephalonia, and Zante [Zakinthos]. The seventh island is Cythera; while historically it groups with the other islands, in that it remained under Venetian rule for an extended period, geographically and linguistically it is intermediate between the Heptanesa, the Peloponnese, and Crete.

Lefkada displays northern Greek vowel reduction, which has been explained (HDMS 540; Ioannis Poulos, 1933) as a borrowing from adjacent Roumeli, where the Lefkadites held their pastures. Following Contossopoulos, I exclude the island of Lefkada from this group, although Triandafyllidis’ (1981 [1938]) isogloss includes it. There has not been an extensive enough study to determine whether Lefkada linguistically patterns with Roumeli or the Heptanesa overall.

⁴¹Contossopoulos’ (1994) bibliography lists only three works on the Peloponnese: Koukoules’ 1908 study on the dialect of northeast Laconia, Litsas’ 1968 study on the phonetics of Trifyllia—

example, a crucial deviation of Corfiot from SMG in the usage of *pu* as a complementiser (§6.8) is ignored in all works on Heptanesian, and is explicitly mentioned only in a parenthetical remark in Pernot's (1934) grammar of Tsakonian.



Map 10. Mainstream Modern Greek Dialect Groups

The major difference between Heptanesian and Peloponnesian lies in language contact; while the Peloponnese was under Turkish rule for some four centuries, and was extensively settled by Albanians in *xv* AD, the Heptanesa spent minimal time under Turkish rule, and were ruled by Venice or local Italian nobles more or less continuously from *xii* to *xviii* AD. As a result, Heptanesian does not display the Turkish lexical influence displayed by the other mainstream Greek dialects, but commensurately displays much greater lexical influence from Italian. There are some differences in phonology and morphology, but they are relatively minor. There is no evidence of Albanian or Slavonic influence on Heptanesian—notwithstanding the fact that Corfu lies only 10 km from the Albanian coast.

2. Northern Greek. Northern dialects are distinctive in that they drop high unstressed vowels and raise mid unstressed vowels. They are spoken north of

and Contossopoulos' (1976 [1975]) own call for more research on Peloponnesian (cf. thirteen works in his bibliography on the Dodecanese.) For the Heptanesa, he mentions only Salvanos' 1916 brief sketch of the dialect of Argyrades in Corfu (little more than a glossary), and Hitiris' 1987 *Corfiot Glossary* (which has some grammatical notes.)

38°N, in general terms⁴² (see isogloss, Map 11). Northern Epirus—that is to say, southern Albania, where a Greek minority remains, speaks a southern dialect; links with Apulia (Höeg 1924) and Mani (Vayacacos 1983 [1978]) have been suggested, but the provenance of the dialect remains obscure.

Newton (1972c:14) notes that Northern dialects have contributed minimally to Standard Greek, and that minority languages spoken in northern Greece (Aroumin, Macedonian Slavonic, Turkish) yield to standard Greek, and not Northern Greek. He further notes that the dialect of Salonica (the major city of northern Greece) differs from Athenian only in vocabulary and morphology (in particular, the northern use of the accusative rather than the genitive for indirect objects: see Map 12)—not in phonology.



Map 11. Northern, Southern, and Semi-Northern dialects, and Albanian-speaking zone in Central Greece (after Triandafyllidis (1981 [1938]:81) and Contossopoulos (1983–1984:163).)

Some northern dialects drop high unstressed vowels, but do not raise mid unstressed vowels; this group includes Eastern Thrace, Mykonos, and Skyros (Newton 1972:182; Contossopoulos 1994:109). Greek linguists class these dialects as Semi-Northern.

⁴²Specifically: in the mainland north of Attica (including Western Thrace, Bithynia, and Eastern Rumelia); northern Euboea; the islands of the north Aegean (Imbros, Tenedos, Lesbos, Lemnos, Samothrace, Thasos, Northern Sporades), part of the Cyclades (northern Andros, Tinos); and Samos (resettled from the north in 1560 (Andriotis 1933:351).)

Morphologically, the Northern dialects are fairly homogeneous; the main differentiation between them, according to Contossopoulos (pers. comm.), is verb endings—whose heterogeneity amongst Greek dialects is prodigious (Newton 1972b; Newton 1975).

The dialect of Constantinople [İstanbul] is a special case within the Northern Greek dialect group. Contossopoulos (1994) does not even consider Constantinopolitan a variant distinct from SMG; it has the northern use of accusative instead of genitive for clitic indirect objects (which also turns up frequently in early Modern texts), but deviates from northern dialects in having a southern (standard Greek) unstressed vowel system.

3. Old Athenian (Mertzios 1964; Thumb 1891). This dialect group includes the old dialect of Athens (now extinct), the dialect of the island of Aegina near Athens (recently extinct), the town of Megara in Attica, and the villages around Kymi in central Euboea. These regions were Greek-speaking islands in a largely Albanian-speaking region (see Map 11). Old Athenian is characterised by phonological archaisms—which gained it the distinction of being singled out as the ‘most barbaric’ of all Greek dialects by several Byzantine and Renaissance Greek scholars.⁴³ In particular, it has [u] rather than [i] as the reflex of Ancient /y/; this is a phonological change known to have taken place in Greek by *x* AD, when a provincial priest is recorded to have been mocked for his modernising pronunciation of /y/ in Constantinople (Browning 1983:57). Furthermore, ‘underlying’ stressed /i/ fails to become an unstressed glide before a vowel: thus, the reflex of Ancient /pe¹dia/ in Old Athenian is /pe¹ðia/, whereas in the rest of the mainstream Greek-speaking world (with the exception of Zante) it is /pe¹ðja/ or a reflex of /pe¹ðja/ (e.g. /pe¹ðja/ or /pe¹θca/).

However, the Albanian populations which cut Old Athenian off from Peloponnesian and Northern Greek are known to have arrived in the region only in *xv* AD, whereas /y/ > [i] dates from *x* AD. At present, we do not know how to explain the archaism of Old Athenian during those five centuries.⁴⁴

Newton includes the dialect of Mani [Maina] (Mirambel 1929) in the southern Peloponnese with Old Athenian, since they share a number of phonological features; Minas (1990), on the other hand, argues for a kinship between this and the other archaic dialect of the Peloponnese, Tsakonian. Cythera, to the south of Mani, is intermediate between Cretan and Old Athenian: it does not form glides, but it has [i] as the reflex of Ancient /y/, and it has palatoalveolar allophones of velars (like Cretan.)

⁴³These include Michael Acominatus, Theodore Zygomalas, and Symeon Cabasilas. An example of Zygomalas’ comments: “And what is worst, were you ever to hear the once most wise Athenians, you would be filled with tears; for just as once they had in abundance the pure and genuine language of the Greeks, so now the barbaric language has spread and is heard above all [By his examples, Zygomalas is clearly talking about old Athenian, not Albanian] [...] The most barbaric dialect is now that of Athens [...] When we hear them, we understand somewhat what they say [...] and we cannot but laugh listening to them.” (cited in Thavoris 1971:16–17)

⁴⁴Slavic incursions are a possibility which I am not aware has been explored yet.

4. **Cretan.** Cretan proper has already been discussed; Newton (1972c) considers most of the Cyclades, and the island of Icaria, to speak a dialect close enough to Cretan that he considers them together. (Other linguists class Icaria as a South-Eastern dialect.) The Cretan influence on the Cyclades can be accounted for both by geographical proximity and colonisation; indeed, the village of Apiranthos in Naxos (whose dialect has been extensively studied—Zevgolis 1956) is originally a Cretan colony, and this is strongly reflected in the local dialect. According to Contossopoulos (1994:46), Patmos in the Dodecanese resembles Eastern Cretan, and was probably resettled from there.

The local dialect died out in the Cyclades very early: in 1918–1919, Voyatzidis (1923:150) could find no speakers whose dialect was unaffected by SMG in Tinos and Mykonos, and only one in Andros and two in Syros. This is because the Cyclades were included in the original Greek state. Indeed, it was the fact that pure SMG, with no local linguistic features, was spoken in Ermoupolis (a city in Syros established by refugees during the Greek War of Independence) that led Voyatzidis (1922) to conclude that SMG was a dialect koine in origin.

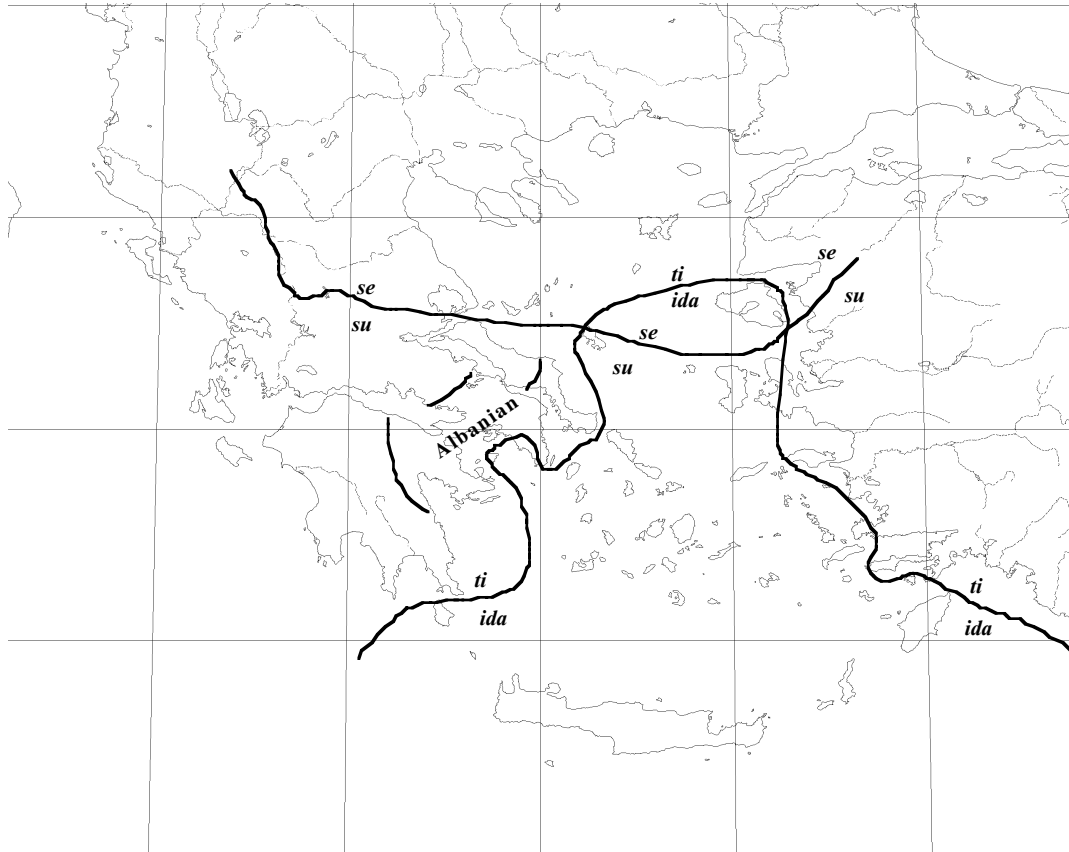
5. **South-Eastern Greek.** Apart from Cypriot (already discussed), this group includes dialects spoken in Chios, the Dodecanese, Amorgos in the Cyclades and the islands between Amorgos and Naxos (Schinoussa, Iraklia), and formerly in Ionia (south-western Turkey). The features Newton (in his phonologically-oriented work) gives to distinguish between the groups are given in Table 24; overall, South-Eastern Greek is phonologically archaic.

Greek dialects fall into a number of isoglosses; while outlier dialects can be fit into at least some of these isoglosses, they are usually left out of such consideration. Traditionally (after Hatzidakis), the most important isogloss is that distinguishing between northern and southern dialects; northern dialects have only three vowels in unstressed position, while southern dialects (and CSMG) retain all five. But the importance of this isogloss is more psychological than linguistic: it affects virtually every word of the language, while the effects of other isoglosses are much scarcer textually. Overall, however, the linguistic difference between Northern and Peloponnesian-Heptanesian is slight. Three major isoglosses distinguishing between Greek dialects—reduced vowel system, accusative versus genitive indirect objects, and *ida* versus *ti* for ‘what’—are given in Maps 11 and 12.

The dialect division invoked most frequently in this work has been formulated by Contossopoulos (1983–1984), who—following on from suggestions by earlier linguists like Dawkins and Newton—has argued that the most important isogloss is that of the word for ‘what?’: *ti* is used on the Greek mainland and the Heptanesa, while *ida*⁴⁵ is used in the Aegean islands (including Crete and Cyprus, but excluding the northernmost Aegean islands), the Pontus, and Old Athenian. This dialect split largely corresponds to the main cultural division amongst Greeks, between mainlanders (στερισιώτες) and islanders (νησιώτες), re-

⁴⁵For more on the derivation of *ida*, see Nicholas (in prep.)

flected in such features as dress (kilts versus breeches), music (clarinet or recorder versus *lyra* [knee-fiddle]), and song (blank verse versus rhymed couplets). Since ‘Islander’ is somewhat vague in its designation, I use Triandafyllidis’ terms **Western** and **Eastern Greek**—taking the opportunity to group Old Athenian and Cretan with South-Eastern Greek, and clearing up their inconsistent grouping by Triandafyllidis as Western dialects.



Map 12. Characteristic isoglosses of Greek: genitive (*su*) vs. accusative (*se*) for dative, and *ida* vs. *ti* for ‘what’ (after Triandafyllidis (1981 [1938]:81) and Contossopoulos (1983–1984:164).)

As Contossopoulos argues, the split also correlates (roughly) with other linguistic features, such as lexical diversity (Contossopoulos 1983), clitic positioning, /ʏ/-epenthesis in *-evo* verbs, and palatoalveolar allophones. Furthermore, while Peloponnesian-Heptanesian is phonologically relatively more archaic than either Northern or Eastern Greek (although South-Eastern Greek retains gemination), in other respects Eastern Greek is typically more archaic, particularly with regards to verb morphology and lexicon.⁴⁶

⁴⁶Cf. Dawkins (1940:13): “It is odd that it is exactly in this most archaic region [Eastern Greek, and South-Eastern Greek in particular] that the language shows the most widely reaching phonetic changes, changes so striking that the words are often unrecognizable until one knows the rules of the game.”

A.3. *The Balkan Sprachbund*

The Balkan Sprachbund is the first instance of extensive areal linguistic contact to have been documented (Sandfeld 1930),⁴⁷ which has made it well-known in general linguistics. It includes as core members:

- **Albanian**—the southern **Tosk** dialect to a greater extent than the northern **Geg** dialect. Tosk, on which contemporary standard Albanian is based, also includes the variants **Arberësh**, spoken in Southern Italy, and **Arvanitika**, spoken in Greece. The Arvanites⁴⁸ emigrated to central Greece in *xiv*–*xv* AD; Albanian was spoken in a coherent zone in Central Greece,⁴⁹ and also in scattered settlements in Western Greece.⁵⁰

The Arberësh emigrated to Italy from Albania⁵¹ and later on Greece, in *xv*–*xviii* AD, although there were already Albanian-speakers in Southern Italy in 1272 (Çabej 1994 [1976]:85). Arberësh is now spoken by some 200,000 people; unlike Italo-Greek, it is not spoken in a small number of enclaves, but is scattered throughout the southern half of Italy.

- **Bulgarian**, the standard variant of which is based on the dialect of the Varna-Tărnovo region, in the country's northeast.
- **Macedonian Slavonic** (often grouped together with Bulgarian by linguists as the dialect chain **Bulgaro-Macedonian**.) Standard Macedonian is based on the dialect of the area of Veles, Prilep, and Bitola, in the region's west; the last of these is some 15 km from the present-day Greek border (although closer to 60 km from the Slavonic/Greek linguistic boundary). Thus, Standard Bulgarian and Standard Macedonian Slavonic represent opposing ends of the Bulgaro-Macedonian linguistic continuum.
- **Eastern Romance**—including **Romanian (Daco-Romanian)**, spoken in Romania; **Megleno-Romanian**, spoken on the border between Greece and the Republic of Macedonia; **Istro-**

⁴⁷For a history of Balkan linguistics, see Schaller (1975:37–48). Although such nineteenth-century linguists as Schleicher and Miklosich had already noticed the affinity of the Balkan languages, Sandfeld's is the first systematic study.

⁴⁸The Albanian-speakers of central Greece are usually named in the linguistic literature by their Greek name, *Arvanites*, and their language *Arvanitika*. Their autonym is *Arbëror*, and their name for their language, *Arbërisht*.

⁴⁹Boeotia, Attica, Southern Euboea, Northern Andros, Corinthia, Argolis.

⁵⁰Epirus, Aetolia & Acarnania, Achaëa, Elis and Laconia.

⁵¹According to Çabej (1994 [1976]:90), "Italian Albanian represents an early stage of Tosk, leading back to dispersal at the end of the Middle Ages. [...] This linguistic branch, detached some centuries ago from the common trunk of the language [...] presents common phonetic, morphological, syntactic and lexical traits with, on the one hand, the Albanian of Greece, separated in an even more remote period, and on the other with the language of the older authors of Northern Albania, in *xvi* and *xvii* AD [...] In its constituent traits, Italian Albanian, like Greek Albanian, belongs to Southern Tosk, or rather South-Western Tosk, and in an ethnolinguistic sense to the dialects of Labëria [extreme southern Albania] and Çamëria [Greek Thesprotia] in the broad sense of those terms."

Romanian, spoken in Dalmatia; and **Aroumin** or **Macedo-Romanian**, spoken in the southern Balkans, including Greece).⁵²

Serbo-Croatian is considered a peripheral member of the Sprachbund, although the Štokavian⁵³ dialects, and the southeastern (Torlak) Štokavian dialects in particular,⁵⁴ are more closely involved in it.

The relative status of **Greek** in the Sprachbund has been subject to debate. Schaller (1975:37) argues it belongs in the ‘second degree’ of Balkan languages with Serbo-Croatian, but Feuillet (1986:22) rejects this (Greek contains nine of the eleven ‘fundamental’ Balkanisms by his count, while Serbo-Croatian contains only two), and includes Greek as a core member.⁵⁵

The similarity between the grammatical systems of the Balkan languages is striking, and the Sprachbund must be considered as a causal factor in the diachrony of any language spoken in the area—including Greek. Constraints of scope (and my lack of fluency in Albanian, Macedonian, Bulgarian, Rumanian and Serbo-Croatian)⁵⁶ force me to omit this component of the research. But as shown by investigations I have conducted (Nicholas 1998a), the story of *pu* is closely bound to that of its Balkan counterparts, particularly Serbian and Macedonian Slavonic *što* and Bulgarian *deto*. The difficulty lies in identifying

⁵²The nationalistically-motivated debate on whether the speakers of Aroumin are ‘latinised Hellenes’ (linguistically assimilated by the Romans) or settlers from Romania has yet to be resolved satisfactorily (Lazarou 1986 [1976]). The question of whether Daco-Romanian and the other Romanian language variants constitute the same language or not has similarly political ramifications; in this study, they are considered separate languages.

Aroumin or *Armuni* is the autonym of Aroumin-speakers; as an ethnicity, they are known in the wider literature by their Greek name, *Vlachs*. The Aroumins of Serbia are called *Cincars*; *Macedo-Romanians* is a Romanian appellation. Winnifrith (1987:7) estimates there to be 50,000 Aroumin-speakers, of which 30,000 live in Greece. At the turn of the century, as Winnifrith discusses, their numbers must have been closer to half a million.

⁵³*Štokavian* is one of the three major dialects of Serbo-Croatian; these dialects are named after the word for ‘what’ in each dialect—*što*, *ča*, and *kaj*. Štokavian is spoken in Serbia, Bosnia-Herzegovina, and eastern Croatia; it includes both standard Serbian (*Ekavian*) and standard Croatian (*Jekavian*)—the two are distinguished by their reflex of Old Church Slavonic *ě*. (A third variant, *Ikavian* (spoken in Bosnia and Dalmatia) does not have literary status.) *Čakavian* and *Kajkavian* are spoken in western Croatia; Kajkavian is a transitional form approaching Slovenian, while Čakavian is geographically restricted to Istria and Dalmatia.

⁵⁴Torlak Serbo-Croatian is spoken in southern Serbia and the northernmost part of the Republic of Macedonia. The question of whether Torlak is part of Serbo-Croatian at all, as opposed to the Bulgaro-Macedonian dialect continuum, has nationalist ramifications—as do so many other such questions in the Balkans. For an overview of the issue, see Birnbaum (1966).

⁵⁵The role of Romany (Gypsy) has not been considered in any work on Balkan linguistics; for my own part, I can only take on face value Hatzisavidis’ (1993:589) comment that “In syntax there are no particular influences from Greek to observe”—i.e. Greek and Romany syntax do not display such syntactic similarities as to evidence extensive language contact, either uni- or bi-directionally. Indeed, such lexical influence as Romany has exerted on Greek has been mediated from marginal sociolects—urban slang and homosexual cant (*kaliarda*); the low prestige of Romany-speakers in Greece is sufficient explanation of this.

⁵⁶As Andriotis & Kourmoulis (1968:21) point out, the Balkanological work on particular languages tends not to be done by native-speaker linguists (“those who are by definition the most qualified to attempt it”), for political reasons. To be fair, Andriotis’ (1992 [1960]) own contribution to Macedonian Slavonic linguistics is not innocent of nationalist preoccupations.

the origin of Balkanisms. Substratum accounts, alluding to hypothetical features of extinct and largely unattested ancient languages such as Thracian or Illyrian, are occasionally made; since they cannot be disconfirmed, however, they are judged by most scholars to be unscientific, and are rejected.

The influence of Greek, as the prestige language of the region in the Middle Ages,⁵⁷ must have been considerable, and both Sandfeld and Greek linguists such as Andriotis attribute most Balkanisms to Greek influence. For some Balkanisms, such as infinitive loss, a Greek origin is widely accepted, due to morphological evidence (Joseph 1981). But Greek does not share all Balkanisms; and Feuillet (1986:26) thinks it unlikely that Greek could have exerted such a profound influence without completely wiping out the other Balkan languages.⁵⁸ The same holds for the earlier dominion of Latin, with the proviso that the Slavs did not arrive in the region until Greek had started displacing Latin as the administrative language of the Byzantine Empire. And indeed, Latin did come close to 'wiping out' Albanian, whose modern vocabulary is substantially Romance.

The fact that Bulgaro-Macedonian, geographically at the centre of the Sprachbund, is the one language group to possess all the fundamental Balkanisms lends support to a 'melting-pot' or adstratal theory, whereby the Balkan languages converged; the role of bilinguals would have been decisive in this respect, just as it is known to have been decisive in other Sprachbunds.⁵⁹ So a non-Greek origin is entirely possible for some Balkanisms shared by Greek.

The time frame at which Balkanisms arose is not clear, even if one rules out the substrate theory. There was extensive Slavic settlement in the Greek peninsula in *vi–viii* AD; yet Old Church Slavonic, known to have been the Slavonic dialect spoken near Salonica in *ix* AD, does not display many Balkanisms. Although Balkanisms need not have all arisen at the same time (Sandfeld 1930:215), it seems probable that most arose between *x* AD and *xv* AD, by which time all Balkan languages are attested in their modern forms, and the modern ethnic distribution of the Balkans was established. Significant Albanian, Bulgaro-Macedonian and Aroumin-speaking (Vlach) populations coexisted with Greek speakers on the Greek mainland until this century; the Bulgaro-Macedonian and Vlach populations were concentrated in Northern Greece, but

⁵⁷In addition, Greek was the prestige language of Romania during Ottoman rule; it was the language of the ruling class, which was Phanariot (aristocratic Greeks from Istanbul).

⁵⁸Sandfeld (1930:178) found it improbable that a language with the prestige of Greek could have been influenced by "a barbarian language unsupported by any political authority." Galton (1981:255) mischievously comments that "if one looks more closely at the Greek history of the era [*ix* AD]—namely, when the Empire was besieged by enemies along almost the full extent of its frontiers, and while oriental-inspired Iconoclasm [religious dispute] had not yet been silenced—one must come to the conclusion that the Byzantine 'political authority' had more pressing business to take care of than the preservation of the infinitive".

⁵⁹There are also other contacts worth noting; in particular, there is significant affinity between Albanian and Romanian, strongly suggesting erstwhile cohabitation (Rosetti 1968:33; Sandfeld 1930:145).

the Albanian-speaking zone, as discussed, was in the centre of the Greek peninsula.⁶⁰

The importance of non-standard dialects in contact linguistics is crucial, as is the fact that language contact is implemented by bilinguals. These factors have not been traditionally given due attention in studies of Balkan linguistics, as Joseph (1992b:130) notes, discussing an early attempt to attribute features of Bulgarian to Rumanian influence although there is no evidence Rumanian was a prestige language to Bulgarian-speakers:

One general consequence of accounts such as Weigand's and others like it has been an overemphasis on comparisons of standard languages rather than regional dialects, even though the contact between individuals, in certain parts of the Balkans at least, more typically involved nonstandard dialects.

A.4. Further dialect scope

There are some regions of the Greek-speaking world underrepresented in my corpus; for future researchers, I outline where the lacunae felt the most are situated. The dialects of Greek are now dead or moribund, so to remedy such lacunae, it is more profitable to look in past recorded text, than to collect new text at this date—particularly when syntactic phenomena of the type considered in this work are amongst the first to perish in linguistic standardisation. In most instances, the past text is probably already available. Time restrictions meant that I limited my research time in Athens to the Centre for the Compilation of the Historical Dictionary of Modern Greek (Κέντρον Συντάξεως του Ιστορικού Λεξικού της Νέας Ελληνικής), which is the main dialectological clearing-house of Greek; but there are certainly plentiful dialect texts to be had in the Centre for Folklore Research (Κέντρον Ερεύνης Λαογραφίας) at the Academy of Athens, the Folklore Libraries (Σπουδαστήρια Λαογραφίας) of Athens University and Aristotle University (Salonica), and the Centre for Asia Minor Studies (Κέντρον Μικρασιατικών Σπουδών) in Athens.

Outlier dialects have been treated well by researchers, precisely because they are outliers: for all but one, I have had all the text and grammatical information I could need, although I would have liked the extent of coverage of Apulian in my corpus to have been closer to that for Calabrian,⁶¹ and it is regrettable that not all villages of Cappadocia are represented by published grammars—a gap

⁶⁰Arvanitika and Aroumin are now dying out. (For a detailed account of the linguistic attrition Arvanitika is undergoing, see Tsitspis (1981).) The Bulgarian population of Greece was subjected to population exchanges with the Greek minority in Bulgaria. (The distinction between Bulgarians and 'Slavophone Hellenes' was necessarily dictated by the borders resulting from the Balkan Wars: thus, the Kukuš (Kilkis) Slavs were expelled, while the Lerin (Florina) Slavs remained in place.) The Greek minority in Albania likewise remains *in situ*. For obvious reasons, the use of Macedonian Slavonic was not looked upon favourably in post-war Greece; up to 100,000 speakers remain (estimates vary and are unreliable), but Macedonian Slavonic, too, is now subject to linguistic attrition. (It is perhaps indicative that Christos Sidiropoulos, a prominent activist for Macedonian minority rights in Greece, is reported not to be fluent in Macedonian Slavonic.)

⁶¹I obtained a copy of Morosi (1870) in October 1998—too late for inclusion in my survey.

particularly acute for Delmeso. Here and there, there are gaps which can no longer be filled: the coverage for Nikopolis is slight, villages in Calabria like Cardeto and Amendolea ceased speaking Greek early this century; relative to the other outliers, Tsakonian is underrepresented. The major gap in outlier dialects, however, is Mariupolitan, severely underrepresented in both my corpus and my bibliography. Although the dialect is now dying out, this is a gap that something can still be done about, with the extensive corpus of '30s literary Mariupolitan (resumed in the '80s), and the texts recorded by Ukrainian linguists from the '50s on. I have not had access to this text for this research.⁶²

Coverage of mainstream Greek is far better with Eastern than Western Greek—for the simple reason that Western Greek, being the group that includes CMSG, has been seen as less interesting linguistically. Indeed, the indifference of Greek linguists until fairly recently towards recording texts⁶³ or syntax (or for that matter what traditional grammar counted as syntax—which includes the present subject matter) has made the linguist concerned with such matters far more grateful to the Greek folklorist, who has rendered a great service in consistently recording the language of texts accurately. As the examples used in this study show, the mainstay of my corpus has been the mainstay of the folklorist's: folktales.

Folklore studies do not place as great a premium on geographical coverage as does dialectology; and while HDMS coverage is broad, there are still regional gaps. This has been reflected in my coverage. I have needed much more Macedonian material than I have had access to, particularly from the area connecting Central Macedonia to Eastern Thrace—Drama, Kavalla, Thasos, and the Western Thracian littoral.⁶⁴ Considering the vitality of Cretan and Cypriot, I have found surprisingly little material in those dialects, although I doubt I have missed any major features. Old Athenian is *terra incognita*, and I regret I have been unable to locate a copy of Mariana Kambouroglou's *Παραμύθια* ('Fairy Tales', 1912), in which some Old Athenian colouring might have still been extant.⁶⁵ And I would have welcomed information from the other western edge of

⁶²Rather late in the piece (August 1998), I have gained access to two recent anthologies of Mariupolitan poetry and prose (Kiryakov 1988; Kiryakov 1989). This has been too late for me to integrate into my research, particularly as no translation of the Mariupolitan has been provided, my reading knowledge of Pontic is of only limited help in deciphering the text, and I do not yet have access to either a grammar or a dictionary (if such exists!) of the dialect.

Cultural exchanges between Greece and Mariupol have started in the '90s, and the Greek government has announced plans to send Standard Greek teachers to the region. The prospect of Mariupolitan yielding to Standard Greek instead of Tatar, Russian or Ukrainian is no comfort to the linguist.

⁶³Most of the manuscripts at the Historical Dictionary archives are glossaries; with a few commendable exceptions (such as Manesis' work on Marmara and Zevgolis' on Apiranthos), extensive transcribed discourse does not appear, and extensive texts do not appear routinely in manuscripts until the '60s.

⁶⁴I did not have the opportunity to use the folkloric resources available in Salonica.

⁶⁵Panagis Skouzes' memoirs are of interest, since he was an Athenian native before Athens was settled as the capital of the Greek state. I have the text, but have not had time to go through it for this study.

Thracian, Imbros, Tenedos, and the Dardanelles; HDMS 280 was compiled in Gallipoli, but does not have enough text to say much about complementation, and Greek linguists have not had ready access to Imbros and Tenedos since the islands were incorporated into the Republic of Turkey. The ethnic Greek populations of the islands are now dying out.

There are two major mainstream dialects I have minimal information on. One is Maniot; Mirambel's (1929) grammar stops at morphology, and I have not found extensive texts in the dialect.⁶⁶ Blanken's (1951) grammar of Corsican Maniot (now extinct) is somewhat better, although I have not had access to its texts volume.

The other, on which my knowledge is minimal, and which is not covered by any HDMS material, may also be Maniot: this is the dialect of Northern Epirus—or to use less irredentist terminology, the language of the ethnic Greeks of modern southern Albania, around Himara (Himarë).⁶⁷ What we do know about Northern Epirot is that it is a southern dialect—far north of the reduced unstressed vowel system isogloss—with archaic disyllabic *-'ea*.⁶⁸ This, plus the fact that a variant of Albanian linguistically contiguous with Tosk was spoken south of Northern Epirus, in Thesprotia (Çamëria), means Northern Epirot cannot be indigenous to the region. There is a local tradition that Himara was colonised from Mani (Dendias 1926:75); Vayacacos (1983 [1978]), a Maniot linguist who has collected some data from Himariots in Corfu, does not rule out this possibility.⁶⁹

Without any texts, I cannot make any judgement myself; for now, I know nothing of Northern Epirot syntax, and cannot say how it fits into the diatopic account given.⁷⁰

⁶⁶Too late for this survey, I have obtained a copy of Kassis (1983), a collection of Maniot folktales.

⁶⁷Anagnostopoulos (1925) also mentions as southern idioms the Greek spoken in Argyrokasto (Gjirokastër) and Delvino (Delvinë) in Southern Albania, the Dryinopolis (Buliarata) district (?), Pogoni (near Delvinaki, Ioannina, Greece; presumably, modern Pogoniani, a mere 2 km from the present-day Albanian border), and Paramythia (Thesprotia, Greece). The reason for the use of southern vocalism in these regions is no clearer than that for Himara. Anagnostopoulos claims that the Himariot dialect, like Maniot, was “more conservative and purer, spoken by populations which lived in virtual autonomy during Turkish rule, particularly because of the nature and geography of their regions, and that the dialects in questions, particularly that of Himara, separated from their surrounding dialects and underwent slower evolution than them, preserving a more archaic and more faithful picture of the mediaeval Greek vernacular.” Why Himara or Argyrokastro would have been significantly more archaic than any number of mountains in Northern Greece (not counting those where Aroumin is spoken!) is not clear.

⁶⁸Vayacacos (1983 [1978]:9) reports that the two villages next to Himara still speaking Greek, Drymades and Palasa, speak a semi-northern dialect; he attributes this to influence from adjoining northern Greek dialects.

⁶⁹Höeg (1924:293), the only linguistic analysis of the dialect I have found, proposes that Himara was colonised by Apulian Italiots after the Turkish raid on Otranto in 1480—a position vigorously refuted by Dendias (1926), who establishes that Otranto has never been a Greek town.

⁷⁰It may soon be too late to gather dialectal data from the region: not only has the ethnic Greek minority looked to Greece—and Standard Greek—for much of recent history, but the current economic dependence of Southern Albania on Greece, driving Tosk, Northern Epirot (ethnic

Greek) and Farshariot (Albanian Vlach) alike to learn CSMG, is confounding the linguistic layout of the region even further.