1. INTRODUCTION

Ο Νηρέας έκανε το κορόιδο.
΄Αλλαξε φάτσες, έγινε έτσι, έγινε αλλιώς κι' όλο μουρμούριζε:
—Δεν είμαι γω ο Νηρέας.
Ο Ηρακλής δεν το 'φαγε.
—Λέγε, ή μένεις δεμένος σαν μοταρντέλα.
Είδε κι' απόειδε ο Νηρέας, το φανέρωσε το μυστικό του περιβολιού.
Nereus played dumb.
He changed his face, he turned into this, that, and the other, and kept muttering.
—I ain't Nereus.
Hercules didn't fall for it.
—Start talking, or you're staying tied up like a salami.
Nereus saw this wasn't getting him anywhere, so he revealed the secret of the garden to him.
(TsifM 445)

The transition from Classical to Modern Greek was marked by a variety of grammatical changes. These changes included several grammaticalisations,[1] giving rise to salient function words used in Modern Greek. Indeed, it can be argued that a major component of the typological difference between Classical and Modern Greek lies precisely in the presence of these function words. The most noteworthy of these grammaticalisations are:

- *hína→ *hina→ na:*
  'in order to' (Classical)→
  subjunctive marker (well under way in New Testament Greek)
  (Burguière 1960; Joseph 1983b; Joseph 1990);
- *hópou(→ ˈopu)→ opu→ pu:*
  Relative adverb 'where' (pre-Classical)→
  relativiser (by 500 AD) (Bakker 1974)→
  particle introducing (typically factive) complement and adjunct clauses;[2]
- *oudèn→ ðen:*
  'nothing'→
  realis negator 'not' (by *ix* AD)
  (Jannaris 1897; Zampelios 1986 [1852]);

[1]Grammaticalisation theory is discussed in §2.
[2]The particle *pu* is cognate to the interrogative ˈ*pu* (Ancient Greek *poû*) 'where?', but is not directly derived from it; this is a misconception which appears frequently in the literature (e.g. Joseph & Philippaki-Warburton 1987:166). See §5 for further discussion.

- *t^{h}élo: hína→ θelo na→ θena→ θa:*
  'I want to'→
  future tense particle (*ix–xv* AD)
  (Jannaris 1897; Kriaras 1969–1993).

Of these grammaticalised particles, *na* and *pu* in particular have become functionally generalised to such an extent in Modern Greek that, almost every time a subordinate clause needs to be attached to a matrix, either *na* or *pu* is employed. Frequently both particles turn up, either in competition or in complementary distribution, as §3 illustrates; and both particles have a wide range of semantic functions, leading a steady stream of Greek linguists from Koraïs in the early nineteenth century onwards to refer to them as Protean—alluding to the mythical sea god who kept changing shape.[3]

The story of the grammaticalisation of *pu* has not been as fully elaborated as that of *na* and *θa*.[4] The grammaticalisation of *pu* up to the point it became a relativiser (~500 AD) is described in Bakker (1974); its subsequent development, however, was ignored by the authoritative Greek diachronicists of last century, Jannaris (1897) and Hatzidakis (1975 [1892]), and has only been taken up in the last two decades.

Recent work on *pu* has been primarily motivated by Greek linguists' attempts to provide an account for the Modern Greek complementiser system; the synchronic literature on the subject has been extensive (Christidis 1981; Christidis 1982; Christidis 1983; Delveroudi 1994; Delveroudi, Tsamadou & Vassilaki 1993a; 1993b; 1994 [1993]; Ginzburg & Kolliakou 1997 [1995]; Kakouriotis 1982; Papadopoulou 1994a; Svalberg 1992; Vande Ostinje 1985; Varlokosta 1994a; 1994b; 1994c; 1995 [1993]), but no clear consensus has emerged yet on the semantic conditioning involved. Only two researchers, Tasos Christidis and Iris Papadopoulou, have attempted to bring diachronic factors to bear in accounting for the distribution of *pu*. Both consider *pu* in the context of the linguistic systems they function in—particularly its functional competition with *na*. However, both accounts have shortcomings.

Christidis (1981; 1982; 1983; 1986) uses grammaticalisation theory to account for the distribution of *pu*, and of *pu*$_{\text{COMP}}$ (*pu* as complementiser) in particular. However, he does not consider the diachronic career of the particle in his account. Instead, he appeals to the property of grammaticalisation known as *persistence* (§2.2.2), whereby the etymology of a grammaticalised word constrains its subsequent grammatical functions. Relying on this property, Christidis considers it sufficient to adduce the etymology of *pu* (<*hópou*, a stationary locative adverb) in order to account for the subsequent semantic functions the particle

[3]Nereus in this chapter's epigraph was another sea god who shared Proteus' propensity for shape-shifting.

[4]Meillet (1921 [1912]) gives *θa* as an example of grammaticalisation in his definitional paper on the subject; *na* is covered extensively in Haspelmath (1989), and Joseph (1981).

took on. This means that Christidis examines only the endpoints of the grammaticalisation of *[o]pu*, but not the interval in between.

Papadopoulou (1994a; 1994b) has a more comprehensive framework of grammaticalisation theory in place, but her account is essentially an expansion of Christidis', and displays a similar attitude to diachronic data. No instances are given of the reanalysis of *pu* in historical texts, and the use of grammaticalisation theory is limited to providing a constraint on the synchronic use of pu_{COMP} in Contemporary Standard Modern Greek (CSMG).

Several issues arise out of the focus of Christidis' and Papadopoulou's research; in giving a comprehensive account of the grammaticalisation of *pu*, these issues need to be addressed:

1 Both researchers emphasise result over process in grammaticalisation. The choice between result and process reflects a larger controversy in grammaticalisation theory (§2.4), over whether metaphor or metonymy has explanatory primacy. That is to say, does grammaticalisation work by transferring words to new conceptual domains *(metaphoric extension),* or by the reinterpretation (reanalysis) of words in ambiguous contexts *(metonymic extension)?* If the former is the case, then pointing out the endpoints of the grammaticalisation makes explicit the conceptual domain transfer involved, and is explanatorily adequate. But if the latter is the case, then what matters are the contexts where the grammaticalising particle is used, permitting its semantic change; the endpoints are not as important as the intervening process.
   Papadopoulou acknowledges this dichotomy in grammaticalisation theory, but argues that the two alternatives are equivalent, with metaphor the synchronic result of metonymy. I contend that metaphoric extension is at best an epiphenomenon of metonymic extension in this particular case, and that metonymy can cope with unevennesses in the data which metaphor, in its propensity to over-generalise, would not consider.
2 Neither researcher pays any attention to the diatopic issues involved in the distribution of *pu;* indeed, no diatopic survey of the distribution of *pu* has hitherto been prepared. Quite apart from the descriptive importance of such a survey, diatopic diversity is inherently diachronic (differentiation in space is often associated with a lag in time), and the diatopic distribution of *pu* raises a number of problems for Christidis' and Papadopoulou's accounts—particularly in privileging a Peloponnesian-based dialect koine, which has become the standard language of modern Greece merely by historical accident.
3 With the exception of a few papers by Christidis and Tzartzanos' *Syntax* (1991 [1946, 1963]), there has been no systematic attempt at a treatment of the third major function of *pu*, other than as relativiser and complementiser—namely, its function in introducing various sentential adjuncts, either alone or in collocations. While this function of *pu* has been integrated into Christidis' and Papadopoulou's diachronic accounts, there is no synchronic description of the conditioning factors on connective *pu* to compare to the ex-

tensive (if inconclusive) work on pu_{COMP}. The prodigious paradigmatic spread of pu_{ADJ} is an excellent illustration of how pervasive a grammaticalisation can become in a language, and I essay a description of the distribution of *pu* in the totality of its functions.

4 The accounts on the synchrony and diachrony of *pu* have hitherto been based on native-speaker intuition, rather than corpora; indeed, in the absence of any computerised Modern Greek corpora until recently, Modern Greek corpus linguistics is still in its infancy. Corpora become particularly useful with respect to the distribution of pu_{COMP}, since native-speaker intuitions on the subject appear to admit some variability. A characteristic of this thesis is its reliance on corpora where available.

5 Minimal attention has been paid to the external factors in the history of Greek which may have affected the distribution of *pu*: diglossia, dialect koineisation, and language contact. Regrettably, constraints in time and research scope prevent me from going into areal factors.[5] The koineisation which gave rise to CSMG does not seem to have had a significant impact in determining the distribution of *pu*; diglossia probably has done so, but it is difficult in Modern Greek to extricate diglossia from register and the spread of mass literacy. There is brief discussion in this thesis on these issues with regard to corpora representing particular registers, but a proper sociolinguistic account also has to await further research.

To summarise: in this study, I provide an account of the fortunes of the particle *pu* since 500 AD. I concentrate on its non-relativiser functions, which constitute the grammaticalisation of the particle subsequent to its initial grammaticalisation as a relativiser. This study aims to be more comprehensive than previous studies; I therefore consider several issues which have not been looked at closely to date—including the distribution of *pu* in dialect, its use to introduce adjuncts, and the insights that can be brought to the analysis of *pu* from corpus linguistics. With regard to general historical linguistics, I attempt to pinpoint some shortcomings of the metaphor approach to grammaticalisation as applied by Christidis and Papadopoulou, and to demonstrate the value of looking at linguistic systems comprehensively, in diatopy, diachrony, and functional competition. (Regrettably, scope does not allow me to investigate the mediaeval career of *pu* directly, but the modern diatopy is itself quite illuminating as to its history.)

In §2, I introduce the core concepts and issues of grammaticalisation theory, which forms the theoretical framework on which my analysis is based. In §3, I give a listing of the various functions *pu* has in Standard Greek, as well as detailing the extant synchronic accounts of the distribution of *pu*. In §4, I discuss in some more detail the semantics of the Greek complementiser paradigm, in which *pu* is the most difficult element to describe. In §5, I give the existing diachronic accounts of *pu*, and describe the diachronic background to the gram-

[5]My preliminary investigation (Nicholas 1998a) suggests the areal factor is significant.

maticalisation—including comparable function words in Classical and Middle Greek.

The chapters up to this point serve as background to the corpus data investigated in §6 (complementiser *pu*) and §7 (connective *pu*). This is accompanied by an excursus, contrasting *pu* with words of similar functionality but different etymology in various dialects, to disprove the privilege of localist etymology, and detailing the allolexes of *pu* (§B). The data considered is synthesised into a coherent account of *pu* in §8. In §A, I give a historical background of Middle and Modern Greek and its dialects, introducing the diachronic and diatopic distinctions made in the body of the thesis.[6]

1.1. Methodology

In contrast to the previous work on *pu,* my work concentrates on the diachrony and diatopy of the particle. As a result, this is a corpus-driven work. For some texts, I have drawn up an exhaustive concordance of all instances of *pu;* for others, I have merely dipped in to the texts for representative instances.[7]

Exhaustive concordances have been drawn up for the following texts (with word-counts in the thousands):

1 Kostas Tahtsis: *The Third Wedding* (1963; 118). This is a novel which (sensationally for its time) is written in a variant of Greek admitting more elements of Puristic than had been usual in literature, in an attempt to reflect more faithfully the language used by the Athenian bourgeoisie. It has been proposed that Tahtsis was self-conscious in going about this (Kazazis 1981)—linguistic self-consciousness being hard to avoid in a language as institutionally diglossic as Greek.[8]

2 Makriyannis' *Memoirs* (1829–1851; 150). Makriyannis was an illiterate soldier in the Greek War of Independence, who after the war learnt enough writing to be able to write his memoirs. His writing is adulated by Greek scholars as a model of demotic writing; although his morphology does display definite signs of influence from the Puristic official language, his syntax is indeed quite oral.[9]

---

[6]This dissertation does not cover all there is to say about the development of *pu*. I have already completed in draft form treatments of the areal status of *pu,* comparing it to similar Balkan developments (Nicholas 1998a), the use of *pu* in collocation in various Greek dialects (Nicholas 1998b), and etymologies of connectives similar to *pu* in outlier dialects, Tsakonian *p^{h}i* (Nicholas 1998f) and Pontic *pi* (Nicholas 1998d).

[7]For the diachronic and diatopic background against which these texts were selected, see §A.

[8]I have had access to soft-copy of this text, for which my thanks to John Burke.

[9]"Makriyannis's language is the vigorous common speech of the peasantry of central Greece, enriched with a host of expressive words and phrases borrowed over the course of centuries from Turkish and Italian, and vivified by the great treasury of Greek folksong [...] It is, one might say, a 'nonconformist' Greek, a language untainted by the syntactic contortionism and lexical necrophily that was to become the bane of the Greek bureaucratic 'establishment' [...] Anger and excitement sometimes make Makriyannis inarticulate. He jumps from a historic past to the present of a running commentary more often than an English reader finds comfortable; this has been smoothed down. He may set out on a long sentence, and, half way over, find him-

3 Jean Psichari: *Historical and Linguistic Matters* (1886; 25) and *My Voyage* (1888, 1905; 65). Psichari was an expatriate Greek linguist of Constantinopolitan descent, and was the first major modern advocate of demotic Greek. His language is a Neogrammarian demotic, eschewing Puristic elements in favour of dialectal constructions and coinages which the demoticist establishment subsequently repudiated. Psichari is another valuable witness of pre-literary Greek, although it is an eclectic Greek, stamped with the author's sense of mission.[10]
4 All texts available to me in the outlier dialects: Silliot, Western Cappadocian, Pharasiot, Pontic, Mariupolitan, Livisiot, Tsakonian, Apulian Italiot, and Calabrian Italiot (although I have relied on the partial concordance of TNC rather than go through TNC exhaustively myself.) The texts are listed in §C.
5 *Honour and Money* (1912; 20) and *Hamlet* by Konstantinos Theotokas, as exemplars of Corfiot.
6 The poems of Constantine Cavafy, as exemplars of Constantinopolitan.

Of the works listed, Tahtsis' has been taken as representative of CSMG, and more realistically representative than most literary CSMG works, which avoid Puristic elements. Makriyannis' has been used as a model of pre-literary vernacular Greek syntax; being from Roumeli, Makriyannis' native dialect is close to Peloponnesian, so his language usage is as close as we can get to proto-SMG. These two texts have thus been used as exemplars of the pre-literary and post-literary standard Greek language—endpoints of its development, through which a global picture of the linguistic variant can be formed.

Other texts have been selected for dialectological reasons. The outliers are investigated exhaustively; the relative paucity of published outlier dialect texts, and the importance of these dialects in representing extreme developments of Greek, make this approach a necessity. I also exhaustively look at Corfiot and Constantinopolitan, in an attempt to formulate quantitative descriptions of their aberrant complementation paradigms (§6.1, §6.8), which have not been the subject of such study before.

I have made more eclectic use of the following:

1 Works by Nikos Tsiforos. Tsiforos, a humorist, used Greek slang extensively in his works written in the 1960s. The informal register of his work is an ex-

---

self lost in the syntactical undergrowth, draw breath, and hack his way through." (Makriyannis 1966 [1850]:xx–xxi)

[10]"From the beginning, he faced a dilemma: on the one hand he felt the need to present the nation with a written language based on the common elements of the dialects, and on the other his ardent desire to become the unique creator of the new national language led him to impose a specific version of Greek which felt like an organic part of his experience and being. This desire is responsible for the great number of non-common forms found in the first edition of *My Voyage*. He was proud to be a Constantinopolitan, and reminds us that Constantinople 'is a significant centre [...] indeed, it was a centre before Athens was.'" (Mackridge 1988:41–42)

cellent source of some of the more colloquial constructions, such as irrealis-*pu θa* and bare-*pu* (§3.8).

2 *Hellas-L. Hellas-L* is an electronic mailing list, in which Greek-speakers exchange email of a mostly informal nature, in romanised Greek. The participants are typically university-educated first generation Greek-speakers, either studying or working abroad; the number of participants in Greece is increasing, but even these are usually educated overseas.
There are several difficulties with the corpus. First, it is Internet-based discourse, and subject to the quirks of that genre, which have drawn a good deal of discussion; it lies at an uneasy balance between oral and written discourse, and has idiosyncracies particular to it (Malone 1995). Second (as the examples cited show), there is a good deal of code-mixing with English involved—although, I would suggest, not significantly more than is usual in the oral discourse of contemporary educated Greeks, and unlikely to range beyond lexis to syntax. Finally, the *ad hoc* romanisations used make searching the texts very difficult; for instance, I have decided to use only `pou,` and not `poy,` `nou,` or `noy,` as the usual romanisation of που (*pu*) which I search for.[11]
Notwithstanding, *Hellas-L* is a voluminous corpus of contemporary colloquial Greek (I estimate some 8.5 million words in my disk corpus, excluding quotations),[12] and I use it as a source on the distribution of constructions in CSMG. I have the archives from 7 November 1996 to 31 January 1998 on disk (95 Megabytes), and have performed most of my searches on this corpus; I have supplemented this with searches on earlier archives (ranging back to March 1994) through the website `www.dejanews.com`. (The mailing list itself has been in existence since 1988.)

3 Archival dialectal material from the *Centre for the Compilation of the Historical Dictionary of Modern Greek* (Κέντρον Συντάξεως του Ιστορικού Λεξικού της Νέας Ελληνικής) at the Academy of Athens. I have on computer some 2300 entries for *pu* and other Modern Greek complementisers. Of these, 1500 are copied from index cards which are to be eventually incorporated into the *Historical Dictionary of Modern Greek;* these index cards, in turn, are derived from fieldwork notebooks stored at the Historical Dictionary Centre. The remaining entries I have culled from those same notebooks. This material has very good geographic coverage of mainstream Greek.

4 Texts in individual dialects depending on availability, as set out in each section.

There has been much Greek dialectal research: the archives of the Historical Dictionary Centre alone have over 1200 manuscripts—mostly fieldwork notes by

---

[11]In March 1997 there were in the *Hellas-L* corpus (as a raw count, including cited discourse) 6728 instances of the string `' pou '`, 1730 of `' poy '`, 280 of `' nou '`, and 0 of `' noy '` (excluding instances of *nu* 'mind').

[12]There are 14.5 million words in the corpus, but 416 out for the 990 instances of `' kai pou '` in the corpus are cited from other participants; I have extrapolated accordingly.

the Centre's dialectologists, ranging back to last century. But little work has been done to integrate this research into a coherent body: Newton's (1972c) comprehensive generative treatment of dialect phonology is yet to be matched in morphology or syntax. At any rate, most of these 1200 manuscripts concentrate exclusively on lexicon and phonetics; extended texts were only recorded regularly from the '60s onwards. Greece still does not have a linguistic atlas: Newton's (1972a) isogloss maps of Cyprus and Contossopoulos' (1988) dialect atlas of Crete have not been matched for any other Greek-speaking region (despite Contossopoulos' persistent lobbying). Sixty years on, the Historical Dictionary has covered only three of the twenty-four letters of the Greek alphabet, in five volumes. And there are still sizeable gaps in our knowledge of Greek dialects—particularly those closest to CSMG. Regrettably, little can be done now to add to our knowledge of Greek dialect; researchers simply have to make do with what has been preserved.