Oxford University Working Papers
in Linguistics, Philology & Phonetics

Volume 7

May 2002

Edited by

Ina J. Hartmann

Andreas Willi

Contents

Editorial note

Oxford University Working Papers in Linguistics, Philology & Phonetics presents research being undertaken by past and present staff and graduate students from the University of Oxford. The current volume concentrates on topics in Comparative Philology and Historical Linguistics. The photograph on page x was taken by Francesca Schironi.

Comments on the papers included here are welcome: the authors' and editors' addresses are listed on the following pages. To get in touch with the editors or for further information you can also write to:

The Centre for Linguistics and Philology
Clarendon Press Institute
Walton Street
Oxford, OX1 2HG
United Kingdom

The current volume of Oxford University Papers in Linguistics, Philology & Phonetics will soon be available on the web at the following URL:

http://www.clp.ox.ac.uk/publications/owp

Ina J. Hartmann   Andreas Willi

Contributors' addresses

Richard Ashdowne
New College
Oxford OX1 3BN
richard.ashdowne@new.ox.ac.uk

Mario Cortina Borja
Institute of Child Health
Centre for Paediatric Epidemiology
    & Biostatistics
University College London
30 Guilford Street
London WC1N 1EH
M.Cortina@ich.ucl.ac.uk

Wolfgang de Melo
Corpus Christi College
Oxford OX1 4JF
wolfgang.demelo@ccc.ox.ac.uk

Diana Gibson
Christ Church College
Oxford OX1 1DP
diana.gibson@chch.ox.ac.uk

Ela Harrison
Department of Classics
University of California, Berkeley
7233 Dwinelle Hall #2520
Berkeley, CA 94720-2520
U.S.A.
elahw@stanford.edu

Ina J. Hartmann
St. Hugh's College
Oxford OX2 6LE
ina.hartmann@st-hughs.ox.ac.uk

Catherine Mary MacRobert
Lady Margaret Hall
Oxford OX2 6QA
catherine.macrobert@lmh.ox.ac.uk

Anna Morpurgo Davies
Somerville College
Oxford OX2 6HD
anna.davies@somerville.ox.ac.uk

Oxford English Dictionary
Oxford University Press
Great Clarendon Street
Oxford OX2 6DP
oed3@oup.co.uk

John H. W. Penney
Wolfson College
Oxford OX2 6UD
john.penney@wolfson.ox.ac.uk

Philomen Probert
Wolfson College
Oxford OX2 6UD
philomen.probert@wolfson.ox.ac.uk

Elizabeth Tucker
The Oriental Institute
Pusey Lane
Oxford OX1 2LE
elizabeth.tucker@orinst.ox.ac.uk

Andreas Willi
Seminar für Klassische Philologie
Universität Basel
Nadelberg 6
CH- 4051 Basel
Andreas Willi@unibas.ch

Jason Zerdin
Lexicon of Greek Personal Names
Rooms 312 & 314
Clarendon Building
Bodleian Library
Oxford OX1 3BG
jason.zerdin@classics.ox.ac.uk

About the contributors

Mary MacRobert, Anna Morpurgo Davies, John Penney, Philomen Probert, and Elizabeth Tucker all teach at Oxford. Jason Zerdin took his doctorate at Oxford in 2000 and is now working for the *Lexicon of Greek Personal Names*; Andreas Willi completed his Oxford D.Phil. in 2001 and is currently at Basel as Post-doctoral Research Assistant; Ela Harrison followed her Oxford BA in Classics with an M.St. in linguistics before moving to Stanford. From 2002-3 she will be a graduate student in the Department of Classics at the University of California, Berkeley. Richard Ashdowne is currently completing his M.Phil. at Oxford, while Wolfgang de Melo, Diana Gibson, and Ina Hartmann are working at their Oxford doctorates. T. S. Dhesi was a student of Mario Cortina Borja in the Department of Statistics at Oxford; he is now at Fitzwilliam College, Cambridge, while Mario is at the Centre for Paediatric Epidemiology and Biostatistics of University College, London. Philip Durkin is Principal Etymologist for the *Oxford English Dictionary*; the other contributors to the *OED* article, Anthony Esposito, Neil Fulton, Jane McCauley, Samantha Schad, Tania Styles, and Joanna Tulloch, all belong to the etymology team which works for Oxford University Press literally across the road from the Oxford Centre for Linguistics and Philology.

Preface

Anna Morpurgo Davies

On a Tuesday in January 2002 students and teachers in comparative philology congregated as usual in Somerville College for a bread-and-cheese lunch. They were surprised to see on the table a particularly impressive cake, complete with candles. The cake-makers, who included one of the editors of this volume, explained that their *opus* was meant to celebrate the thirtieth anniversary of the so-called philological lunches which, as they had discovered, had started in January 1972. These lunches normally prelude to a graduate seminar and there is a close link between the Tuesday seminars and the papers collected in this volume, since a number of them were either presented at the seminar or prompted by work done for it. Volume 7 of *Oxford University Working Papers in Linguistics, Philology and Phonetics* (*OUWPLPP*) now provides a more permanent celebration of that thirtieth anniversary and of the long sequence of Tuesday seminars in comparative philology.

OUWPLPP 7 concentrates on comparative philology, i.e. on comparative and historical linguistics, just as its predecessor *OUWPLPP* 3 (1998). The whole series of working papers in linguistics, philology and phonetics was started seven years ago at the prompting of James Higginbotham, Professor of General Linguistics. On that occasion Dr Coleman, the Director of the Phonetics Laboratory, pointed out in the preface to vol. 1 that:

> General linguistics and phonetics are enjoying something of a revival in Oxford at the moment; comparative philology has long been a research strength.

It is indeed true that Oxford has a long tradition in this field — a tradition which goes back to a much earlier period than 1972. The chair of comparative philology was first founded for Friedrich Max Müller in 1868 and was the first chair that the University established at its own expense without being prompted by a benefaction. The first graduate degree in the subject (the Diploma in Comparative Philology, which eventually turned into the current M.Phil. in General Linguistics and Comparative Philology) was established in the late 1920s at a time when in Oxford there were few postgraduate courses, particularly in the Humanities; the first student to obtain the Diploma (1932)

was C. E. Bazell, who later became the second professor of general linguistics at SOAS following Firth. The first Oxford lecturership in general linguistics dates from 1969 and the professorship from 1978, the phonetics laboratory was established in 1980, but it was only in 1996 that the creation of *OUWPLPP* marked the fact that after a number of hesitations and difficulties theoretical linguistics, phonetics, comparative philology, and Romance linguistics had succeeded in combining their resources at the graduate level under the aegis of the Committee for Comparative Philology and General Linguistics, the 'modern' replacement for the older Committee for Comparative Philology established in 1928. Each volume so far has concentrated on a specific field but clear signs of cross-fertilization are visible in all of them.

The papers which the two editors have collected in this volume include work by present and past staff and students of Oxford University. They cover a wide range of subjects and of languages (Vedic, Greek, Latin, Italic, Slavic, etc.). It is pleasing to note that continuity is guaranteed: two of the contributors (John Penney and Elizabeth Tucker) were among the graduate students who participated in the 1972 philological lunches and a third (C. M. Mac Robert) was an undergraduate at the time. The two editors belong to a more recent generation: one is currently working at her Oxford dissertation on Greek onomastics; the other finished last summer his Oxford doctorate on the Languages of Aristophanes and is currently teaching in Basel. We are all very grateful to them for the immense amount of work they put into this volume and the way in which they cheerfully confronted all sorts of problems, from the most scholarly to the most mundane. Last but not least, they had to compound differences which others might have found insurmountable: one is a PC user and the other a Mac person.

Among the papers collected here only two cannot be immediately labelled in terms of successive generations of philology students and teachers. T. S. Dhesi was M. Cortina-Borja's MSc student in the Oxford Department of Statistics when, like one of our first students, D. A. Ringe of the University of Pennsylvania, he became interested in statistical methods for the genetic classification of languages. Here student and teacher present some of the first results of their collaboration.

The *Oxford English Dictionary* is the most important work of lexicography in the English language and is too well known to require an introduction. But not everybody knows that a monumental work of revision of the original edition is currently going on and that the first results are already available on line. The Oxford philologists and the group of so-called OED 'etymologists', based at

Oxford University Press, have enjoyed close contacts for a number of years and at least two of the etymologists (Neil Fulton and Samantha Schad) were Oxford 'philologists' before they (literally) crossed the road and started to work for OUP. The group summarizes here some of their results and gives a more detailed account of their method than is possible in the concise form of a dictionary item. On an earlier occasion some OED etymologists took the graduate seminar step by step through a sample of their work and left it with the impression that lexicography is as risky, as full of pitfalls, and ultimately as exciting and rewarding as the most dangerous of all sports or the most with-it of all academic pursuits.

All the papers published in this volume share an interest in the history and comparison of languages, and of ancient Indo-European languages in particular, but it would be foolish to assume that they have a common theme or even a common method. In all of them, however, we detect the historical linguist's constant complaint about data: there is never enough evidence, we can never appeal to native speakers' intuitions, we do not believe in *argumenta ex silentio* but we have to use them when we deal with ancient languages, etc. Yet, in spite of the difficulties, new results can be achieved.

Jason Zerdin aims at solving the old problem of why the Homeric and Herodotean -σκον verbs appear to have either intensive or iterative meanings and reaches a conclusion based both on typology and on a close analysis of the semantic value of the verbs: -σκον mainly indicates iteration which is realized as repetition in dynamic situations and as continuous action in stative ones.

Diana Gibson discovers that Classical Greek is rich in periphrastic causatives of the type 'I make you do something', even if the standard grammars normally assume that this is a late development; she now begins to sketch a chronology and a rationale for the spreading of the construction in the late fifth and fourth centuries BC.

Wolfgang de Melo produces a close study of Latin sigmatic futures like *faxo*, *faxis* and highlights the contrast between main clauses where only *faxo* appears with a future meaning, often in causative constructions and sometimes with a fossilized quasi-adverbial meaning, and the subordinate clauses where there is a multiplicity of verbs (*faxo*, *capso*, *occupasso*, *liberasso*, etc.) which mostly indicate anteriority and belong to the higher registers. The contrast between Plautus and Terence (the latter has practically abandoned the sigmatic forms except for *faxo* in main clauses) is instructive.

Richard Ashdowne concentrates on the Latin vocative in order to puzzle out whether vocative phrases belong or do not belong to the syntax of the sentence;

in contrast with the traditional view he aims at a view of the vocative as syntactically and pragmatically integrated in the sentence.

Two papers deal with prosodic questions. Mary MacRobert studies the way in which Cyrillic Manuscripts in Serbian Church Slavonic of the 13th-14th centuries use spacing to divide prosodic 'words'. Proclitic and enclitics form a graphic unit with orthotonic items, while the first and second element of a compound are sometimes separated by a space. Spaces then seem to mark accentual units or, as in the compounds, units with primary or secondary stress. The conclusion is that we can define clitics for this period and see how they differ from those reconstructed for Old Church Slavonic; at the same time we see that below the level of the sentence or the phrase the main unit is the prosodic unit in contrast with the modern or not so modern concentration on lexical items. The phenomenon is shared by a number of ancient Indo-European languages and would deserve a general typological analysis.

Philomen Probert is also concerned with clitics but this time in Latin. She provides strong support for the Latin grammarians' hesitant observation that clitics call for an immediately preceding accent. Hence we have penultimate accentuation not only in *doctusque,* which follows the general accentuation rule, but also in *utraque* and *itaque* 'and thus' which contradict it. The contrast between *ítaque* 'therefore' and *itáque* 'and thus' is explained because the former is morphologically opaque and consequently follows the general accentuation rule. Here too the phenomenon is typologically easy to match even if this type of analysis was not available to the grammarians.

Greek is the centre of attention for three other contributions. Ina Hartmann turns to onomastics and starts with a well justified complaint about the incoherence and vagueness of the terminology used to describe various types of names (*Kurznamen*, *Kosenamen*, etc.). She then introduces a synchronic classification of names into monothematic and dithematic with further subdivisions and uses her database of Elean names to test the two contrasting views that all Greek names were originally compounds (dithematic) and that they were originally monothematic. The conclusion is that neither view is supported by her data: the monothematic names must have arisen both as abbreviations of dithematic names and as monothematic forms taken from lexical items.

Ela Harrison exploits the data collected in her Oxford master thesis and analyses the letters written on ostraka in the second century AD at Mons Claudianus, a *castellum* in the Egyptian desert. Even in this remote part of the world the remodelling of the Greek case system follows the usual paths of late Greek, but personal names, clitics, and partitive genitives reveal some special

developments which may give us an insight in the way in which the remodelling occurs. In fixed word order contexts such as those which occur at the beginning of letters it looks as if the morphological marking of case can be ignored since pragmatic considerations are sufficient. It is remarkable that the personal names are far more prone than ordinary lexical items (even in the same contexts) to neglect the 'correct' case marker.

Andreas Willi concentrates on etymology; the immediate problem is whether the two Greek presents ἥκω and ἵκω which both mean 'I come' with a slightly perfective nuance ('I have reached') have two completely different origins or not. The standard application of sound laws leads to positing two different roots, but the semantic identity makes us wish for a different solution. Willi concludes that it is possible to find a common source without reverting to pre-scientific phonology. The general point is important; at the very least one must accept that in a situation of this type extra effort is needed to explore all avenues; as Willi says: in etymology semantics is not or should not be *une quantité négligeable.*

Finally, Sabellic and Vedic. The two papers which deal with these languages share a concern for defining the meaning of words and forms through a close analysis of their context.

John Penney concentrates on the deictic system of Sabellic pronouns and shows that, contrary to the standard view, Umbrian *esto-* and *es(s)o-* do not fill slots equivalent to those of Latin *hic* and *iste*, but are part of a single paradigm for 'this'; the *esto-* forms are reserved for the direct cases (except for the feminine nom. sing.) and the *es(s)-* forms for the indirect ones. The etymology poses more problems but some of these at least may be solved and the whole pronominal system of Sabellic now appears different.

Elizabeth Tucker starts from a rare Vedic noun *r̥gmín-* which is often interpreted as synonymous with *r̥kvan-* 'singer of verses'; she shows that in fact *r̥gmín-* must go back to a root with a voiced labiovelar or velar stop, and derives it from a reconstructed *r̥gma-* which also accounts for the adjective *r̥gmíya-*. A discussion of the forms from the present stem *r̥ñj(a)-* leads to the conclusion that the RV provides evidence for two separate roots. The inherited root *$H_3er\hat{g}$- 'to go straight' accounts for some participles which are active and intransitive, but the finite nasal-infix middle forms which are transitive go back to a different *r̥j-* root which meant something like 'to stimulate, to stir, to excite' and had an original labiovelar or velar. It is striking that this middle present occurs in the same collocations with *girā́* or *gīrbhíḥ* 'song(s)' as the adjective *r̥gmíya-* and the noun *r̥gmín-*. A *r̥gmín-* is a person who stimulates (through songs). All links are contextually assured, and the confusion with *r̥c-* 'hymn, verse' is later. Phonology and morphology fall into place but so does the semantics.

OUWPLPP 3 contained eight papers; *OUWPLPP* 7 contains thirteen papers. We do not anticipate, nor do we wish for, a similar rate of increase in future volumes. Our wish is that the staying power of philological lunches is matched by equal resilience in the Working Papers. It remains to be seen whether any of the papers, past, present, or future, can match the perfection of our celebratory cake.

RV *ṛgmín-*, *ṛgmíya-*, and *ṛñjate*

Elizabeth Tucker

A connection between *ṛgmín-* and the root *ṛc-* 'to praise in verse, to sing praises' is found in the Indian tradition (e.g. Sāyaṇa, RV 1.100.4: *ṛgmy arcanīyaḥ*) and this has served as the basis for its translation in all Western Lexica, e.g. the St. Petersburg Dictionary (Böhtlingk-Roth) 'preisend, jubelnd', and Grassmann (1873) 'singend, jubelnd'. In spite of its unexpected -*g*-, it has been regarded as a synonym of *ṛ́kvan-* 'singer of verses, singing praises' which is likewise confined to the RV but is attested more frequently. However, *ṛgmín-* occurs in only two RV hymns neither of which allow its precise sense to be established from the context:

RV 9.86.46

> *ásarji skambhó divá údyato mádaḥ   pári tridhā́tur bhúvanāni arṣati*
> *aṃśúṃ rihanti matáyaḥ pánipnataṃ   girā́ yádi nirṇíjam ṛgmíṇo yayúḥ*

Translated by Geldner (1951-7: 3.84):

> 'Er ward ausgegossen, der Pfeiler des Himmels, der emporgehaltene Rauschtrank; er fliesst mit drei Bestandteilen um die Welten. Die Dichtungen lecken an dem schreienden Stengel, wenn die *Versdichter* mit ihrer Rede Staat machen.'

Renou (1955-69: 9.36):

> 'Il vient-de-se-déverser, étai du ciel, breuvage offert, il coule autour des mondes, (ce dieu) triparti. Les pensées-poétiques lèchent la tige miraculeuse, quand *les (poètes) pourvus de strophes* sont allés avec le chant vers la robe-d'apparat (qu'a revêtue le *soma*).'

RV 1.100.4

> *só áṅgirobhir áṅgirastamo bhūd   vṛ́ṣā vṛ́ṣabhiḥ sákhibhiḥ sákhā sán*
> *ṛgmíbhir ṛgmī́ gātúbhir jyéṣṭho   marútvān no bhavatu índra ūtī́*

Geldner (1951-7: 1.128):

> 'Er (Indra) ist mit den Angiras' der beste Angiras, mit den Bullen der (beste) Bulle, mit den Freunden ein Freund, *mit den Lobsängern ein Lobsänger*, unter den Wegen der beste - der Indra soll uns in Begleitung der Marut zu Hilfe kommen.'

Renou (1955-69: 17.34):

> 'Lui, il est parmi les Angiras le meilleur des Angiras, le mâle parmi les mâles, étant un ami parmi les amis, un *détenteur-de-strophes* parmi les *détenteurs-de-strophes*, la meilleure (voie) parmi les voies, etc.'

The morphological analysis of this nominal stem also creates difficulties. According to one explanation (Grassmann 1873: 277, Macdonell 1910: 140) it is to be segmented *r̥g-mín-* and represents a secondary derivative in *-mín-* from the root noun *ŕ̥c-* f. 'song of praise, hymn, verse'. In other words, it shows 'compositional sandhi' because it is based on a preexisting nominal stem, whereas *ŕ̥kvan-* is a primary derivative built with the inherited suffix *-van-* (cf. *pī́van-*, *pī́varī*: Greek πίων, πίειρα), and may be compared to other RV stems with nomen agentis value that consist of the accented root + suffix *-van-* (cf. RV *yúdh-van-* 'fighter', *yáj-van-* 'worshipper', *pát-van-* 'flying', etc.: Macdonell 1910: 131-2, Wackernagel-Debrunner (1896-1957: II/2.895-6).

The Old Indic suffixes *-vin* and *-min* are traditionally considered to represent Indic innovations that result from a blend of inherited *-vant* (< IE *-u̯ent-*) and its conditioned variant *-mant*[1] with the semantically proximate suffix *-in*.[2] In Vedic secondary adjectives in *-vant* or *-mant* frequently stand beside other adjectives derived from the same nominal bases by means of the suffix *-ín*, and the meaning of the two sets of derivatives is often similar:[3] cf. e.g. RV *áśvavant-*, *aśvín-* 'possessing horses', *sáhasvant-*, *sahasín-* 'possessing might', *arcivánt-/arcimánt-*, *arcín-* 'possessing rays', AV *kṣīrávant-* 'containing milk' (of pail), *kṣīrín-* 'having milk' (of cow).

---

[1] Agreements between Vedic and Avestan indicate that *-mant* originally occurred after Indo-Iranian stems in *-u-*, *-au-*, *-us-*, and probably arose from *-vant* by dissimilation. In Old Indic it was further extended, e.g. to stems in *-i-* and *-iṣ-*: cf. Wackernagel-Debrunner (1896-1957: II/2.880-3).

[2] Wackernagel-Debrunner (1896-1957: II/2.918-19); cf. Bloomfield (1910-11: 51): '-*vín* and -*mín* stems are in the main *-vant* and *-mant* stems modulated over into *-ín* stems'. But no attempt has been made to explain in detail how the blend is supposed to have come about.

[3] Wackernagel-Debrunner (1896-1957: II/2. 332-41). *-ín* predominates for the description of animate beings. The origins of the Old Indic suffix *-ín-* and the problems raised by the pattern of inflection shown by this class of secondary nominal derivatives will not be discussed here.

From the earliest examples *-vín* does indeed appear to function as a variant of *-vant*, but at first its distribution is very restricted. In the RV we only find formations in *-ā-vín* (e.g. *māyāvín-* 'possessing magical power' from *māyā́-* f. 'magical power, power of illusion', *dvayāvín-* 'characterized by duplicity' from *dvayá-* n. 'duplicity') and *-as-vín* (e.g. *rakṣasvín-* 'harmful' from *rákṣas-* n. 'harm', *tarasvín-* 'endowed with superiority' from *táras-* n. 'might, superiority').[4] It is only in the case of the derivatives from *s*-stems that parallel formations in *-as-vant-* and *-as-vín-* coexist (e.g. RV *námasvant-*, *namasvín-* 'reverent' from *námas-* n. 'reverence', AV *yáśasvant-*, *yaśasvín-* 'possessing glory' from *yáśas-* n. 'glory').[5]

The case of *-mín* is not at all similar, as some of the handbooks observe (Macdonell 1910: 140, Witney 1889: 472-3). Examples where *-mín* coexists with *-mant* (e.g. *go-min-* 'possessing cows' beside earlier *gó-mant-*) may be quoted from Epic Sanskrit (Wackernagel-Debrunner 1896-1957: II/2.776), but such parallel formations are lacking in Vedic. The RV stems in *-mín* have mostly arisen in one of the following two ways. Firstly, they are formed with the suffix *-ín* on the basis of primary nominal stems in *-má-* or *-ma-*, e.g. *dhūmín-* 'smoky' from *dhūmá-* m. 'smoke', *śuṣmín-* 'vehement' from *śúṣma-* m. 'vehemence' (cf. adj. *śúṣma-* 'vehement'), *bhāmín-* 'radiant' from *bhā́ma-* m. 'radiance, brightness' (Wackernagel-Debrunner 1896-1957: II/2.776, §622a). Even if the primary nominal stem is originally an adjective in *-má-* (e.g. *ruk-má-* 'shining' < *luk-mó-*), there is always clear synchronic evidence for its use as a substantive (*rukmá-* m. 'jewel', *rukmín-* 'shining, jewelled'). Hence this group merely shows the same plentifully attested Old Indic process of deriving a secondary adjective in *-ín* meaning 'possessing X' from a substantive with thematic stem as *áśva-* : *aśvín-*, *dyumná-* : *dyumnín-*, etc. Secondly, in a smaller number of cases *-mín* clearly results from the recharacterization of stems in *-mí-* with a further suffix *-n-*: RV 9.98.6 *ūrmíṇam* 'wave', cf. *ūrmí-* m. 'wave', RV 8.6.12 *tuvikūrmin* (voc.), cf. *tuvikūrmí-* (epithet of Indra) 'going about strongly (?)' (Wackernagel-Debrunner 1896-1957: II/2. 776, §622b). Here we have an

---

[4] RV 10.108.5 *áyudhvī*, which was explained by Grassmann (1873: 98) as the nominative of an adjective in *-vín*, represents *áyuddhvī*, an absolutive in *-tvī* (cf. Wackernagel-Debrunner 1896-1957: II/2. 654). RV 5.52.2 *dhr̥ṣadvínas* is connected not with *dhr̥ṣát-* but with *d(h)r̥ṣád-* according to Insler (1999). The frequency of *-as-vín-* must have been a factor in the creation of the hapax legomenon *śatasvín-* (in a triṣṭubh cadence, RV 7.58.4 *vípro ... śatasvī́*) 'possessing a hundred', in addition to the analogy of *sahasrín-* 'possessing a thousand'.
[5] Once attested *māyā́vant-* in *māyā́vān ábrahmā dásyur* (RV 4.16.9) was probably backformed for metrical reasons from the more frequent *māyāvín*. Otherwise there are no pairs in *-ā-vín/-ā-vant*. Cf. Narten (1980) on the case of *āmayāvin-* 'possessing pain', where a process of reanalysis led to the backformation of a noun *āmaya-* 'illness'.

identical process of recharacterization to the one that occurs occasionally in the RV for other types of stems in *-í* which are employed as appellatives: e.g. RV 1.85.3 *abhimātínam* 'opponent', cf. *abhimātí-* m. 'opponent', RV 5.52.12 *kīríṇas* 'of the singer', cf. *kīrí-* m. 'singer, poet'.

Apart from *r̥gmín-* the only other case in the Saṃhitās where a full-blown suffix *-mín-* has been suspected is *iṣmín-*, an epithet of the Maruts (RV 1.87.6, 5.87.5, 7.56.11) and Rudra (RV 5.52.16). Bloomfield (1910-11: 49-51) explained *iṣmín-* < **iṣu-mín* 'possessing arrows, armed with arrows', a parallel formation to *íṣumant-* (RV 2.42.2, of the Maruts 5.57.2). Even though this meaning was accepted by Neisser, Geldner, and Renou, the assumed loss of *-u-* is unsupported, since the only parallel to which Bloomfield could point is the loss of -u- before the 1st pl. inflections *-mas*, *-ma*, *-mahe*, etc., in *-nu-* presents. The loss here is a morphological phenomenon, probably of Indo-Iranian date (cf. Cowgill 1968: 264), and the impossibility of taking it as a regular phonological development is demonstrated by *íṣumant-* itself (which occurs in *sūktas* attributed to the same poet Śyāvāśva Ātreya). The traditional meaning 'impelling, driving, stormy' is not impossible in the contexts (cf. Mayrhofer 1986-96: 1.202), hence some connection with the root *íṣ-* 'to send, to impel' is probably to be sought. Conceivably, it might represent a secondary derivative in *-mín-* from the root noun *íṣ-* f. 'strength, refreshment, invigourating drink, libation', but, if so, it shows regular sandhi. However, it is more likely that *-mín-* arose in one of the ways outlined above which have parallels in the RV: either *iṣm-ín*, a secondary derivative in *-ín-* from an unattested action noun **iṣma-* (but Old Avestan *aēṣəma-*, YAv. *aēṣma-* 'Fury, Wrath' point rather to an Indo-Iranian noun **ais-ma-* with full-grade root), or *iṣmí-n-*, an *-n-* extension of **iṣmí-*. It is clear that *iṣmín-* presents so many uncertainties of morphology and meaning that it cannot help to explain *r̥gmín-*. We shall return to it only at the end of this paper.

The suffixes *-mín-*, *-vín-* are employed to form a few derivatives from root nouns in later Vedic: for *-mín-* the earliest example is *vāgmín-* (Ś.Br. 10.3.3.1) based on *vā́c-* f. 'speech'; for *-vín-*, *vāgvín-* (AVŚ 5.20.11), supposedly based on the same root noun (but see below), and *sragvín-* (JB 2.103) from *sráj-* f. 'garland'. However, the RV itself contains further evidence which suggests that to interpret *r̥gmín-* as such a derivative from *r̥c-* is anachronistic.

Firstly, in order to explain why we find *r̥gmín-* rather than **r̥gvín-*, it would be necessary to suppose that *-mín* here replaced *-vín*, just as *-mant* begins in the RV to replace *-vant* after stems in *-r̥* ('new' *tvā́ṣṭr̥mant-* 6.52.11, *hótr̥mant-* 10.41.21 versus 'old' and frequent *nr̥vant-*). However, *vipŕ̥kvant-* (RV 5.2.3)

shows that after *ŕk-* (belonging to the root syllable) *-vant* was possible[6] and therefore **ṛk-vín-* or **ṛg-vín* would also be expected.

Secondly, even if (in spite of the above arguments) we postulate a derivative in *-mín* it is still difficult to explain why we find *ṛgmín-* and not **ṛkmín-*. 'Compositional sandhi' for secondary derivatives in *-vant* and *-mant* only becomes regular in Epic Sanskrit for some categories, such as derivatives from *s*-stems, e.g. *tapovant-* vs. AV *tápasvant-*, *tapasvín-*; cf. also *ātmavant-* vs. RV, AV, TS *ātmanvánt-*, Ś.Br. *ātmanvín-*. But in the RV the regular inherited sandhi is the rule before *-mant*, *-vant*, and *-vín*: e.g. *virúkmant-*, *vipṛ́kvant-*, *vivakvánt-*, *marútvant-*, *námasvant-* (cf. OAv. *nəmax^vant-*), *namasvín-*.[7] RV, AV *dátvant-* (from *dánt-/dát-* 'tooth') vs. AV *pádvant-* (from *pā́d-/pad-* 'foot') shows that monosyllabic bases were treated no differently in the early *Saṃhitās*.

Although it does not appear to have been questioned, it is not certain that the once occurring AV *vāgvín-* is a derivative from *vā́c-* 'voice', or the same word that later appears as *vāgmín-* 'eloquent' (Ś.Br., Manu, etc.; according to Wackernagel-Debrunner 1896-1957: II/2.776, *vāgvín-* > *vāgmín-* by dissimilation). At AVŚ 5.20.11 the War Drum is told: *vāgvíva mántraṃ prá bharasva vā́cam* which Whitney (1905: 256) translates 'Bring forth thy voice as a speaker his discourse'. But it is probable that the simile is meant to emphasize loudness rather than eloquence or mere ability to speak. If the point is that a *vāgvín-* does not mutter a mantra but intones it loudly and clearly, this word could be based on the same root as RV *vagnú-*, *vagvanú-* 'noise'.[8]

For *ṛgmín-* the alternative morphological analysis is *ṛgm-ín-*, a derivative built with suffix *-ín-* from a primary nominal stem **ṛgma-* (Whitney 1889: 472-3, Wackernagel-Debrunner 1896-1957: II/2.776). The attested stem *ṛgma-* (AB) cannot be the base from which *ṛgmín-* is derived as it is an adjective and means '*ṛc*-like, verse-like'. Yet, the existence of an old nomen actionis **ṛgma-* is indicated by the RV adjective *ṛgmíya-* which is traditionally translated 'to be

---

[6] So does the once attested *ŕkvant-* (4.50.5), but as this probably represents a remodelling of *ŕkvan-*, it provides less certain evidence (cf. also *rkvá-* 10.36.5 with accent on suffix).

[7] The one exception is the once attested *pṛ́ṣadvat (barhíṣ-)* RV 7.2.4, which may be explained by analogy with the series of compounds *pṛ́ṣadaśva-*, *pṛ́ṣadyoni-*, etc. This is the type of external sandhi that starts to appear earliest for *–vant* derivatives, cf. e.g. TS *sádvant-*, VS *bṛhadvant-*.

[8] To be compared with Latin *vāgīre* according to Lubotsky (1981: 134), Schrijver (1991: 146). A derivative *vāgv-ín-* from a noun **vāg(u)va-* (cf. *srákva-*, etc.) might be postulated. Cf. AV *śabdín-* 'noisy, roaring' based on *śábda-* m. 'noise'. Alternatively, could *vāgvī́* be a 'loud-crying woman' (f. of an unattested *u*-stem adj. **vāgú- < * ụeH₂g-ú-*)? Cf. RV 10.32.3 *jayā́ pátiṃ vahati vagnúnā sumát | puṃsá íd bhadró vahatúḥ páriṣkṛtaḥ* 'Die Gattin führt den Gatten heim mit schönem Geplauder; des Mannes glückliche Hochzeitsfahrt ist ausgerichtet', and the use of Lat. *vāgīre* for the crying of infants and women.

praised, to be hymned'. In other words, *ṛgma-* : *ṛgmíya-* might be postulated on the basis of the morphological and semantic parallel found in *yajñá-* 'worship': *yajñíya-* 'to be worshipped'. This pair of formations was inherited from Indo-Iranian, cf. Av. *yasna-*, *yesniia-*, and the pattern was extended to produce similar pairs in both branches, e.g. YAv. *vahma-* : *vahmiia-*, RV *stóma-* : *stómia-*, both meaning 'praise : to be praised'. But if behind *ṛgmín-*, *ṛgmíya-* there lies a nomen actionis *ṛgma-*, can the root in question possibly be *ṛc-* 'to sing'? There is no justification at all for assuming 'compositional sandhi' in such a primary formation at any date (cf. *rukmá-* 'jewel', *vákman-* 'speech', etc.),[9] and so the conclusion seems inescapable that here we must have a root ending in a voiced stop *ṛg-/ṛj-*, that is, a root ending in an inherited voiced velar or labiovelar stop (cf. e.g. *yugmá-*, *yugmán-* from *yuj-*, *yunákti*).

From a synchronic point of view three roots with the shape *r(a)j-* may be identified in Vedic:

(1) *raj-* 'to redden' (first attested in the AV: *rajaya* AVŚ 1.23.1, *arajyata* AVŚ 15.8.1), which shows an alternation *j/g* (ppp. *raktá-*, *rāga-* m. 'colour, redness' Br.+), but which it is impossible to connect with *ṛgmín-*, *ṛgmíya-* not only on semantic grounds but also because of the complete absence of zero-grade forms in its derivatives.

(2) *ṛj-/ṛñj-* 'to go straight', which makes only present stems (RV nasal infixed pres. athematic 3rd pl. middle *ṛñjate*, thematic 3rd sg. middle *ṛñjáte*, *-ya-* pres. *abhi ... ṛjyate*, etc.[10]), and is normally connected with the adjective *ṛjú-* 'straight, direct' (cf. Av. *ərəzu-*, Gk. ἀργός, etc.); the verb has accepted cognates in YAv. *rāzayente* 'arrange in line', *ham.rāzayata* 'stretched out', Gk. ὀρέγω 'I extend', Lith. *ręžti* 'stretches', OIr. *rigid* 'stretches out' (Mayrhofer 1986-96: 2.425).

---

[9] The one possible parallel case that has been adduced is the adjective *sagmá-* (RV epithet of Soma, children of Aditi, horses, chariot, etc.) which is traditionally connected with the root *śak-* 'to be able' and translated 'powerful, strong'. But Geldner (1951-7: 1.366 n. 1d) pointed out that in the AV this word belongs to the same semantic sphere as *śivá-*, and that such a meaning is required for *sagména mánasā* (RV 3.31.1). The connection proposed by Bailey (1958: 149-52) with MPers. *sagr*, Zor. Pahl. *sgl* 'content, satisfied' (< Old Iranian **sag-ra-*) seems preferable in spite of the doubts expressed by Mayrhofer (1986-96: 2.604).

[10] *ṛjyantaḥ* (RV 6.37.2, 6.37.3) may not be a *-ya-* present form: it is scanned *ṛjiyantaḥ* at both its occurrences, and has been explained by Hoffmann (*apud* Joachim 1978: 61) as a compound *ṛji-yant-*, containing a 'Caland form' of the adj. *ṛju-*. The stem seen in RV *irajyáti*, *irajyata*, etc., was explained by Peters (1986: 372 n. 26) as a deverbative present in **-ye-/-yo-* based on the athematic nasal present of this root (**H₃ṛng-ye- > iraja-*), but its connections are not at all certain, as in at least RV 1.7.9 the meaning is similar to that of *rā́jati* (cf. Gotō 1987: 271, with references). Also cf. Gotō (1987: 99-102) for arguments against connecting the late Vedic present *árjati*. None of these diachronically obscure formations will be considered further in this paper.

(3) *rāj-* 'to rule, to shine', which is characterized by lengthened grade forms (RV *rā́ṣṭi*, *rā́jati*, etc.). It has often been suggested that *rāj-* continues the same IE root *$H_3erĝ$-* as *ṛj-/ṛñj-*, and that both belong to a large IE family of cognates that comprise not only Gk. ὀρέγω 'I extend', OIr. *rigid* 'stretches out', Lith. *rę́žti* 'stretches', but also Lat. *regō* 'I rule', etc., and the words for 'king' in Old Indic, Latin, etc. Cf. e.g. Benveniste (1969: 2.9-15), Strunk (1987: 385-92), Rix (1998: 270-1).

This larger diachronic grouping has not been universally accepted (cf. e.g. Sihler 1977, Scharfe 1985); but whatever groupings are adopted the cognates outside Indic for (2) and (3) point incontrovertibly to an IE root or roots that ended in a palatal stop. Thus at first sight there seems to be no possibility of connecting *ṛgmín-*, *ṛgmíya-* with any of these familiar Indic verbs.

However, none of the forms that occur for the verb *ṛj-/ṛñj-* (no. 2 above) in fact show whether the root-final *-j-* represents a palatal or (labio)velar stop. For instance, the athematic nasal infixed present, which must undoubtedly represent an ancient formation, is attested unambiguously only by a form where even in the case of a root ending in an inherited plain velar or labiovelar the palatal allophone would be regular: 3rd pl. pres. middle *ṛñjate*. From roots ending in velars which build a similar infixed nasal present cf. 3rd pl. pres. middle *yuñjáte*, but 3rd sg. pres. active *yunákti*, middle *yuṅkté*; 3rd pl. pres. middle *vṛñjáte* but 3rd sg. pres. active *vṛṇákti*, middle *vṛṅkté*. The complete set of athematic forms is unknown[11] because this verb does not survive in Classical Sanskrit or even in late Vedic except in passages based on the RV,[12] and already in the RV a thematicized stem *ṛñjá-* is being generalized. This thematicization must have started from precisely those forms (e.g. 3rd pl. pres. active *ṛñjánti*, 1st sg. pres. middle *ṛñje*) where a root-final palatal would be regular regardless of whether the root originally ended in a palatal stop or not.

In short, the widespread assumption in recent literature (cf. e.g. Werba 1997: 395-6 n. 496) that the root-final *-j-* of *ṛñjate* etc. must represent an inherited palatal stop rests entirely on etymological considerations and the accepted connection with the *ṛjú-* family of words. This in turn is based on the context of some of the RV forms, and particularly on the word-play at RV 4.38.7

---

[11] Böhtlingk-Roth (1855-75: 1.428) posited 3rd sg. middle *ṛṅkté* beside the attested 3rd pl. *ṛñjate*, no doubt because all other attested 3rd pl. middle forms in *-jate* correspond to an attested 3rd sg. middle in *-ṅkte*.

[12] E.g. in the Aitareya Āraṇyaka; the one apparent attestation of an athematic form *ṛñjyāt* in Vedic Prose (JB I 77:7) should be emended to *mṛñjyāt* according to Gotō (1987: 102 n. 48).

túraṃ yatī́ṣu turáyann ṛjipyó   *ádhi bhruvóḥ kirate reṇúm ṛñján*

translated by Haebler (1968: 295) 'unter den schnell gehenden schnell laufend, geradlinig vorschiessend, wirft er bis über die Augenbrauen Staub auf, rasch sich vorwärtsstreckend'. Here *ṛji(pyó) ... ṛñján* appears to represent a similar figure of speech to *túram ... turáyan*.

However, the Western lexicographical treatments of *ṛj-/ṛñj-* have recognized a range of meaning and usage for this verb and its different stems. The St. Petersburg Dictionary (Böhtlingk-Roth 1855-75: 1.428) distinguished the meanings (1) 'sich strecken, ausgreifen', (2) 'erstreben, verlangen nach', (3) partic. *ṛñjasāná-* (a) 'herbeieilend' (b) 'erstrebend'. Grassmann (1873: 280-1), who links *rají-* 'line', superl. *rájiṣṭha-* 'straightest', commented 'Die Grundbedeutung ist: etwas biegsames in eine gerade Linie ausrecken oder ausspannen; *ṛñj-* ... bedeutet daher 1) in gerader Linie *vordringen, vorwärtsschiessen*, ... so auch 2) in der Verbindung mit *yat* (streben) ... 3) ... etwas [A.] *erstreben, hinstreben* zu [A.]'.

As pointed out by Haebler (1968), there are two basic syntactical patterns for *ṛj-/ṛñj-* in the RV. Firstly, intransitive-absolute usage, e.g.:

RV 1.172.2

āré sá vaḥ sudānavo   *máruta ṛñjatī́ śáruḥ*

'Fern bleibe euer *geradlinig vorwärtseilender* Pfeil (ihr gabenschönen Marut)'.

Secondly, a construction with an accusative which Haebler following the 19th-century lexicographers takes as an accusative of goal/movement towards, e.g.:

RV 1.143.7

agníṃ mitráṃ ná samidhāná ṛñjate

'wie zu einem Freunde *eilt geradewegs* zu Agni der Anzündende.'

This construction is often extended by an instrumental (always of a noun meaning 'song, praise', and, here only, also 'oblation'), e.g.:

RV 2.2.5

sá hótā víśvam pári bhūtu adhvaráṃ   tám u havyaír mánuṣa ŗñjate giŕā

'er (Agni) soll als Opferpriester den ganzen Opferdienst besorgen, ihm *eilen* die Menschen mit ihren Opferspenden, mit ihrem Liede *geradewegs zu*'.

Like Grassmann, Haebler takes the basic meaning of the verb to be 'sich in gerade(r) Linie rasch vorwärts bewegen'. This semantic and syntactical analysis has been repeated in many recent treatments (e.g. Haudry 1977: 314-17, Joachim 1978: 60, Gotō 1987: 271, Mayrhofer 1986-96: 2.425). Haudry (cf. de Lamberterie 1990: 1.717) attempted to explain the extended construction (e.g. 4.8.1, 6.15.1 *ŗñjase giŕā*) with accusative and instrumental via an inherited syntactical pattern, quoting e.g. Il. 23.99 ὠρέξατο χερσὶ φίλῃσιν. But there is a considerable difference in sense between 'to hasten towards X with song' and 'to stretch out towards X with one's own hands' (even allowing for figurative employments of a phrase that may originally have described a physical gesture), and the verb stems employed in Homer and the RV are not comparable.[13]

On the other hand, Geldner's RV translation offered a simpler analysis of the second construction, as here the accusative is taken as the object of a transitive verb, e.g. RV 1.143.7 *ghṛtápratīkaṃ va ... agním mitrám ná samidhānā ŗñjate* is translated 'Eurem Agni, dem Schmalzgesicht ... *lässt* der Anzündende *den Vortritt*' and RV 2.2.5 *tám u havyaír mánuṣa ŗñjate giŕā* is translated 'Ihn *zeichnen* die Menschen mit Opferspenden mit Lobrede *aus*'.[14]

---

[13] Opinion is divided about whether the Homeric participle ὀρεγνύς 'stretching out' represents an old nasal present or is an early Greek innovation (cf. Rix 1998: 271 n. 4), but, however that may be, in its two attestations it is followed by the accusative of the word for 'hands' (*Il.* 1.351, 21.37 χεῖρας ὀρεγνύς).

[14] In his RV translation Geldner appears to have taken the verb to mean basically 'to move forward, to press forward' with both an intransitive and a transitive usage, and a semantic development to 'give precedence to, to honour' in connection with the latter use (cf. Geldner 1951-7: 1.124 n. 7b). A similar syntactic analysis was presented in his Vedische Studien (Geldner 1901: 26-35), but here (Geldner 1901: 29) he followed Pischel in rejecting a connection with ὀρέγω, and proposed meanings that were in accord with the Indian tradition: (1) intr. 'glänzen, strahlen, prangen, geschmückt werden', (2) trans. 'glänzend/schmuck machen, schmücken, herausputzen, dekorieren, auszeichnen, feiern'. However, he also distinguished (Geldner 1901: 27-9) a separate root *rj-/sŗj-* 'losstürzen, sprengen, galoppieren' to which many of the intransitive forms were assigned. This suggestion is clearly to be rejected: Santucci (1982: 241-54) has shown how the similarity in meaning results from an approximation in usage that develops in later Vedic, and it was later abandoned by Geldner himself, but it is interesting to note how in his earlier work he saw difficulty in assigning all the RV material to one and the same root.

This interpretation of the syntax in such passages has been adopted by a number of other scholars, who have combined it with a meaning 'to direct' (Macdonell 1916: 374, Lubotsky 1997: 384; cf. Peters 1986: 372 'richten, lenken'), which is closer to that of supposed cognates in other IE languages.[15] But here the comparison encounters morphological problems, since none of the presents with such a meaning are of comparable structure to *r̥ñjate*. In particular the Old Iranian present with this meaning is of a different characterized type, a lengthened grade *-aya-* formation (YAv. *rāzayente*, etc.).[16] Strunk (1987: 390) noted that from a morphological point of view it is difficult to relate the seemingly archaic Old Indic nasal infixed present to the stems in other branches of IE that point to a common inheritance for the IE root *$H_3erĝ$-*. The suggestion by Rix (1998: 271 n. 3), following Peters (1986: 372 n. 26), that the nasal infix created a present stem with transitive value meaning 'to direct, to drive' for the root *$H_3erĝ$-* 'to go straight' is contradicted by the fact that in the RV uncompounded nasal present forms with active inflection are normally intransitive in value (cf. Geldner's note to RV 1.95.7b), e.g. RV 4.38.8 (cf. 4.38.7 and 1.172.2 quoted above):

> *yadā́ sahásram abhí ṣīm áyodhīd   durvártuḥ smā bhavati bhīmá r̥ñján*

'Wenn Tausend ihn (Dadhikrā) bekämpfen, so war er nicht aufzuhalten, wenn der Furchtbare *an die Spitze eilt*.' (Geldner 1951-7: 1.470)

'Quand un millier (d'adversaires) l'a attaqué, il est impossible à retenir; (il est) terrible quand *il s'élance*' (Renou 1955-69: 15.163).

If we want to understand how the forms of this verb are distributed between the two syntactical patterns, it is necessary first of all to exclude from consideration the anomalous 'middle participle' *r̥ñjasāná-* (3: listed under a separate lemma by Lubotsky 1997). Like other RV 'participles' in *-asāná-* it is probably to be connected with an *s*-stem noun and is of denominative origin: cf. *ohasāná-* : *óhas-*, *jrayasāná-* : *jráyas-*, *rabhasāná-* : *rábhas-*, *sahasāná-* : *sáhas-* and AV *bhiyasāná-* : *bhiyás-* (Wackernagel-Debrunner 1896-1957: II/2.236-7, Insler 1968: 16-23). In the case of *r̥ñjasāná-* the *s*-stem *ŕ̥ñjas-* is attested in the RV, but in MIA form, by adverbially employed *áñjas*, *áñjasā* 'straightway' (Geldner

---

[15] Cf. Strunk (1987: 389) who regards the 'primordial meaning' of the IE root *$H_3erĝ$-* as 'to stretch', its secondary meaning as 'to direct, to rule'.

[16] The Av. root *raz-* is glossed 'richten' by Bartholomae (1904: 1514) (cf. Kellens 1995: 82, 'mettre en ligne, tirer un trait droit, aligner, mettre en ordre'), but no other finite forms occur except those built on the pres. stem *rāzaya-*.

1901: 43, Wackernagel-Debrunner 1896-1957: II/2. 230). The root involved is certainly **H₃erĝ-* 'to go directly, to stretch' and because of its formal appearance this 'participle' has been linked to the verb, but its usage can tell us nothing about the usage of the original verbal stems. Once these three 'participle' forms are discounted, a much clearer distribution emerges. Nasal present *middle* forms, both uncompounded and with the preverbs *ni-*, *ā-* are regularly associated with an accusative.[17] So is once occurring *prá ... ṛñjánti* which shows active inflection. On the other hand, nasal present *active* forms without a preverb normally lack such an accusative.[18]

Such a distribution of active and middle forms based on the same stem is hard to understand in terms of the diathesis shown by a single verb. It will be argued here that in fact two verbs are to be distinguished among this set of forms and they have completely unrelated meanings. The active participles reflect inherited **H₃erĝ-*, but the largest group of nasal present forms (both athematic and thematic), which are employed transitively and show middle endings, belong to a root which we will henceforth call *²rj-*. The latter means 'to move emotionally, to stimulate mentally, to stir, to excite', it ends in a (labio-)velar stop, and its allomorph *ṛg-* appears in *ṛgmín-*, *ṛgmíya-*.

Evidence for the meaning of the root *²rj-* comes from Middle Indo-Aryan where according to Haebler (1968) Pali *iñjati*, Buddhist Hybrid Sanskrit *iñjate* continue the Vedic thematic stem *ṛñjá-*, rather than the Vedic verb *iṅg-* 'to move' (RV *iṅgaya*, *iṅgayanti*, cf. *éjati*). Haebler's examination of the contexts shows that this verb is employed to indicate excitement, agitation, perturbation, etc., and the subject is always a person or *kāya* 'body' or *citta* 'mind, heart'. Usually it occurs in negated statements of the sort 'X is not perturbed, X does not care', e.g.

---

[17] Present middle forms: *ṛñjate* 3rd pl. (RV 1.37.3, 1.122.13, 1.141.6, 2.1.8, 2.2.5), *ṛñjase* 1st sg. (4.8.1, 6.15.1, 6.15.4), *ṛñjate* 3rd sg. (1.143.5, 1.143.7), *ā ... ṛñjase* 1st sg. (5.13.6, 10.76.1), *ní ... ṛñje* 1st sg. (3.4.7 = 3.7.8, 4.26.1), *nyṛñjáse* 2nd sg. (8.90.4), *ní ... ṛñjase* 2nd sg. (10.142.2), *nyṛñjáte* 3rd sg. (1.54.2). Active form: *prá ... ṛñjánti* (3.43.6). Exceptions are the athematic 3rd pl. middle form from the once attested compound *sam ... ṛñj-*, which is followed by a locative (RV 1.6.9 *sám asminn ṛñjate gíraḥ*). Possibly also the thematic 3rd sg. middle *ṛñjate* at 5.48.5, which was noted as an exception by Geldner (1951-7: 1.124). *ṛñjata* (RV 5.87.5), probably 3rd pl. middle injunctive (cf. Geldner 1951-7: 2.90, Renou 1955-69: 10.39) rather than 2nd pl. active injunctive/imperative (as Grassmann 1873; 281), is only an apparent exception, as our discussion below will show that an object such as *ródasī* is to be understood.

[18] All participle forms: *ṛñján* (RV 1.95.7, 4.38.7, 4.38.8), *ṛñjatī* (1.172.2). In the problematic verse RV 3.31.1 *ṛñján* may be associated with an accusative, and this will be discussed below.

Nd I 353.31ff.

> *so lābhe pi na iñjati, yase pi na iñjati, ayase pi na iñjati ...*

> 'beim Empfangen erregt er (der Bhikkhu) sich nicht, beim Nichtempfangen erregt er sich nicht, bei Ruhm erregt er sich nicht, bei Nicht-Ruhm erregt er sich nicht ...'

Dhp 81

> *selo yathā ekaghano vātena na samīrati evaṃ nindāpasaṃsāsu na samiñjanti paṇḍitā*

> 'Wie ein Felsblock vom Winde nicht bewegt wird, so sind auch die Weisen bei Tadel und Lob unbeweglich'.

As the above examples show, the simple thematic stem is employed in intransitive function, but in Pali there is also a characterized stem *iñjayati* which has transitive value: Haebler suggests that in MIA the reflex of *ṛñjáte* shows a semantic development resulting from a narrowing of the concrete sense 'bewegt sich rasch in gerade(r) Linie vorwärts' which he postulates for early Vedic. On the other hand the new transitive present *iñjayati* is supposed to preserve the old meaning in phrases such as *lomam pi na iñjayati* 'er sträubt auch nicht ein Haar'. However, these MIA facts could be interpreted differently: the meaning of the root could have remained unchanged and there could merely have been a morphological innovation (of a banal type for MIA) which resulted in the creation of a more clearly characterized transitive stem (*iñjayati*), while the old stem *iñja-* < *ṛñjá-* was confined to intransitive uses. This hypothesis would permit us to seek a similar meaning in Vedic, and to distinguish diachronically the forms with such a meaning from those which mean 'to go straight, to stretch'.[19]

A sense 'to move emotionally, to stimulate mentally, to stir, to excite' would fit all the RV contexts for the simple nasal present middle *ṛñjate*, etc., as here the accusatives invariably designate a god or a human being. In addition, the

---

[19] The differing phonological developments of *ṛja- seen in Vedic *áñjas-*, *áñjas ā* and the Pāli and Buddhist Hybrid Sanskrit verb stems *iñja-*, *iñjaya-* may result from the selection of forms from different Indic dialects. This becomes easier to understand if two diachronically distinct roots are concerned, and makes it unnecessary to exclude the derivation of *áñjas-* < *ṛñjas-* because of the existence of *iñja-* < *ṛñjá-*, as Werba (1992: 17). The entries in Turner (1966: 117-18) show that both developments are represented in Middle Indo-Aryan and New Indo-Aryan.

frequent occurrence of a noun meaning 'song, fine-verses' in the instrumental would fit with such a meaning, since for the Vedic poets their poems are their means of moving the gods to grant their requests. Hence *(agním) ... ṛñjase girā́* could on this basis be translated 'I move/I stimulate Agni with song'. At RV 2.2.5 (see above for full quotation) *tám u havyaír mánuṣa ṛñjate girā́* the humans move Agni by means of their oblations and song, as the following verse requests him to shine wealth on us (*rayím asmā́su dīdihi*). At 1.143.7 *agníṃ mitráṃ ná samidhāná ṛñjate* the kindler stimulates Agni as though he were a friend. Cf. 6.15.1 *imám ū ṣú vo átithim uṣarbúdhaṃ | víśvāsāṃ viśā́m pátim ṛñjase girā́*, translated by Geldner (1951-7: 2.106) 'Diesem eurem frühwachen Gast, dem Fürsten aller Stämme will ich fein mit meiner Lobrede den Vorzug geben'. But it can be suggested that the description of Agni as an early awakening guest has more point if he is being stirred to waken by the poet's song.

Such examples may be multiplied (cf. RV 1.122.13, 1.141.6, 2.1.8, 6.15.4) and in addition it becomes possible to understand the value of the once-attested infinitive form *ṛñjáse* and its parallelism to the infinitive from *stu-* 'to praise'[20] in RV 8.4.17 *vémi tvā pūṣan ṛñjáse | vémi stótava āghṛṇe* 'I seek you, O Pūṣan, for stimulating, I seek you, O glowing one (?), for praising'.[21]

A basic meaning 'to move, to stimulate' also is capable of explaining the value of various *²rj-* compounds that take a direct object. The most frequent

---

[20] Geldner (1951-7: 2.290 n. 17*ab*) regarded the meaning as similar to *ṛñjate girā́* (RV 2.2.5 and 6.15.1). The fact that the '1st sg. middle' form *ṛñjase* is distributed in the same way as a verb of speaking (cf. Schlerath 1984: 206-14) is explained if *ṛñjase girā́ | suvṛktibhis ... ṛñjase* are virtual synonyms for *stuṣe*.

[21] One difficult example requires fuller discussion. In the Atijagatī verse 5.87.5 *sváno ná vo 'mavān rejayad vṛṣā́ tveṣó yayís taviṣá evayā́marut | yénā sáhanta ṛñjáta svárociṣa sthā́raśmāno hiraṇyáhāḥ svāyudhā́sa iṣmíṇaḥ* an overtly expressed object is lacking for *ṛñjáta*, but the same is true for 3rd sg. injunctive active *rejayad* in the main clause. There can be no doubt that *rejayad* is a transitive form because of its morphology and three other RV passages where the same form appears. Renou (1955-69: 10.39) translates 'le taureau (Viṣṇu) ... fait trembler (l'univers)', and RV 7.57.1 *rejáyanti ródasī cid urvī́* (cf. 1.31.3, 1.151.1, 2.11.9 *árejetāṃ ródasī*) points more precisely to an object *ródasī* 'the Two Worlds'. Since 1.54.2 shows *ródasī* as the object of *nyṛñjáse* with another bull (Indra) as the subject, it may be suggested that the same object should be understood for *ṛñjáta* in 5.87.5: the Maruts stir the Two Worlds with the help of Viṣṇu, just as Viṣṇu makes them tremble by making a noise like the Maruts' thunder. Hence perhaps 'Mighty like your thunder, the terrifying rapid strong bull makes tremble (the Two Worlds) — (Viṣṇu with whom is) the swift-going Marut — with whom the powerful ones, having their own light ... stir (the Two Worlds)'. It is possible to interpret *ṛñjáta* as an athematic 3rd pl. middle injunctive form (as Geldner and Renou do) in spite of *vo* in *pāda* a, since this hymn fluctuates between second and third person reference for the Maruts, and the syntax is tortuous because of the insertion of the refrain *evayā́yamarut* in every verse.

ni-ṛñj- (x 7 in RV) was translated by Geldner 'zwingen, nötigen, drängen', but Grassmann and Renou distinguished two meanings (in some cases for the same form).[22] However, if it refers primarily to mental compulsion or perturbation rather than physical knocking down (Grassmann 'jemand [A.] niederstrecken'), it becomes easier to understand why the object is not always an enemy (as 8.90.4 *tvám* [Indra] ... *vṛtrā bhū́ri nyṛñjáse*), but may also be both the Goddesses Rodasī, i.e. the Two Worlds (1.54.2 [Indra] *yó dhṛṣṇúnā śávasā ródasī ubhé vṛ́ṣā vṛṣatvā́ vṛṣabhó nyṛñjáte*), or the Two Divine Hotars (3.4.7 *daívyā hótārā prathamā́ ní ṛñje*) or even all creatures (10.142.2 *agne ... sacíva víśvā bhúvanā ní ṛñjase*). In another passage Agni is causing trouble to trees (1.143.5), *yodhó ná śátrūn sá vánā ní ṛñjate* 'er zwingt die Bäume wie ein Krieger die Feinde' (Geldner 1951-7: 1.201), but here this verb may be employed because of the simile.

A '1st sg. middle' form *ṛñjase* is attested four times for the simple verb, three times in initial verses where its first person reference is beyond doubt, whatever the diachronic origin of such forms. If *ā́ ... ṛñjase* means 'I stretch out towards', this results at RV 5.13.6 *ágne nemír arā́m̐ iva devā́ṃs tvám paribhū́r asi | ā́ rā́dhaś citrám ṛñjase* in the unlikely picture of the poet helping himself to brilliant wealth (Geldner 1951-7: 2.15, 'Agni! Wie die Felge die Speichen so umfängst du die Götter. Ich trage nach ansehnlicher Belohnung Verlangen'); Renou obtained a more normal sense by taking *ṛñjase* as 2nd sg. (with imperative value) 'Dirige vers (moi) un bienfait éclatant', but at the cost of analysing this same form differently in the two passages where it occurs (5.13.6 and 10.76.1: cf. Renou 1955-69: 13.113). If instead we derive this compound from *²ṛj-*, 5.13.6 could be understood as 'O Agni ... I stimulate (you to bring) towards (me) brilliant wealth'. As often in the RV, the preverb *ā́* may indicate an ellipsis of a dependent verb of movement or conveying. Similarly, at RV 10.76.1 the Pressing Stones, who are personified throughout this hymn and elsewhere in Vedic literature, are adressed *ā́ va ṛñjasa ūrjā́ṃ vyùṣṭiṣu* 'I stimulate (you to bring) your energy towards (me) every morning'.

There can be no doubt that *prá ... ṛñjánti* at RV 3.43.6 must somehow mean 'they open'. In the verse *ā́ tvā bṛhánto hárayo yujānā́ arvā́g indra sadhamā́do vahantu | prá yé dvitā́ divá ṛñjánti ā́tāḥ súsammṛṣṭāso vṛṣabhásya mūrā́ḥ* Geldner and Renou both took *ā́tāḥ* as the direct object of the verb in the relative clause ('die abermals die ... Tore des Himmels aufstossen'; 'eux qui dès l'origine assaillent les portiques du ciel'). In the Āprī hymns the 'doors of heaven' are regularly addressed as if they were a deity and implored to open of

---

[22] Renou (1955-69: 10.59) '*ní ṛ(ñ)j-* est ambivalent: "attirer à soi (dans une intention favorable)" 3.4.7, 4.26.1, "attirer pour contraindre (un ennemi)" 1.143.5, 8.90.4; les deux nuances à la fois 1.54.2, 10.142.2 (+ 1.37.3)'.

their own accord (e.g. 1.142.6 *ví śrayantām ... dvā́ro devī́r*), hence a literal sense 'to stimulate forth, to arouse forth' for *pra ṛñj-* would fit this context. Middle inflection is the norm for *²rj-* and its compounds because it is most frequently used in contexts where the worshipper is moving a god for his own advantage (once at 1.122.13 the patrons are moving the poets by means of gifts to perform for the patrons' advantage). In other words, we have the regular Old Indic use of the middle as in *yajate* versus *yajati*. The active inflection of *prá ... ṛñjanti* may therefore be explained because Indra's horses make the doors of heaven open not for their own benefit but to allow Indra passage to the mortal world.

If the compound *sam ... ṛñj-* attested once at RV 1.6.9 *átaḥ parijman ā́ gahi divó vā rocanā́d ádhi | sám asminn ṛñjate gíraḥ* is from the root 'to go straight, to stretch', *asmin* must be explained as a locative of goal following a verb of movement. ('Ihm streben die Lobreden gemeinsam zu', Geldner 1951-7: 1.8; 'Les chants confluent vers ce (dieu)', Renou 1955-69: 17.3). However, since *ṛñjate* represents the same athematic 3rd pl. present middle form which we have argued in all its other five occurrences (RV 1.37.3, 1.122.13, 1.141.6, 2.28, 2.2.5) represents the root *²rj-* and since its subject is *gíraḥ* 'songs', a form of the same noun which occurs in connection with the expression *ṛñjate/ṛñjase gírā*, there is also an argument for identifying this as a compound of *²rj-*. The traditional interpretation has the drawback that a switch from 2nd to 3rd person within a short uncomplicated Gāyatrī verse must be assumed if *asmin* refers to Indra. On the other hand if the locative refers to *divó rocanā́d* in the previous *pāda*, an accusative 2nd person pronoun may be understood from the vocative *parijman* and 2nd sg. imperative *ā́ gahi*, and the force of *sám* could be 'similarly, in the same way'. Hence 'From here (i.e. this world), O circulating one, or from the light heaven, approach. In it (the light heaven) the songs similarly stimulate (you)'. In other words the poets' songs reach Indra even if he is in heaven and stimulate him to come to their offering.

The residue of forms which cannot be assigned on the grounds of meaning and syntax to the root *²rj-* is very small, but since evidence for an exactly matching nasal present from the root **H₃erĝ-* is lacking in other branches of IE, we will make a short digression concerning these forms before returning to *²rj-*, which is our primary concern. There are five active participles, and in four of these cases it is the same form *ṛñján*, occupying the same position at the end of a Triṣṭubh *pāda*. It is not disputed that they mean 'going straight, going directly', nor that the identification of *ṛjipyáḥ ... ṛñján* as a *figura etymologica* at RV 4.38.7 is correct. However, all the attested participles may represent a thematic stem *ṛñjá-*, as the once occurring feminine form *ṛñjatī́* (1.172.2) could be a

parallel case to *siñcatī́ḥ* (RV 10.21.3) beside *āsiñcántīḥ* (RV 5.86.6) from the thematic nasal present *siñcáti*.[23]

The lack of finite forms[24] corresponding to these participles is surprising, but the only finite present form in the RV which unquestionably belongs with the root *$H_3erĝ$-* is *abhí ... ṛjyate* at 1.140.2:

> *abhí dvijánmā trivṛ́d ánnam ṛjyate*   *saṃvatsaré vāvṛdhe jagdhám ī́ púnaḥ*

> 'Der Zweigeborene eilt dreifältig auf die Speise los. In Jahresfrist wächst wieder, was er verzehrt hat.' (Geldner 1951-7: 1.195)

In this verse a preverb/preposition *abhí* indicates the goal towards which the subject of the verb moves, precisely in the manner we should expect for a verb meaning 'to go straight, to stretch'.[25] Moreover, if the RV *-ya-* present seen in *abhí ... ṛjyate* is set beside the Old Iranian lengthened grade *-aya-* present seen in YAv. *rāzaya-*, we have here a typical Indo-Iranian pair of intransitive/ transitive present stems for the root in question (cf. Vedic *yúdhyate/yodháyati*, *búdhyate/bodháyati*, Avestan *būiδyaite/baoδayeiti*).[26] If both Old Iranian

---

[23] Wackernagel-Debrunner (1896-1957: II/2.419). Such feminine participle forms appear in later Sanskrit for Class VI presents (Pāṇini 7.1.80) and their occurrence in the RV is motivated by metre: *máruta ṛñjatī́ śáruḥ* makes a better Gāyatrī *pāda* than *máruta *ṛñjántī́ śáruḥ*.

[24] One finite form — *ṛñjate*, 3rd sg. thematic but with a middle inflection! — cannot be securely attributed to *²ṛj-*, although such an interpretation is not impossible. It occurs in an obscure verse of a hymn to Agni, 5.48.5 *sá jihváyā cáturanīka ṛñjate | cā́ru vásāno váruṇo yátann arím* 'That four-faced (god, i.e. Agni) ?stimulates/goes directly? with his tongue, (like?) Varuna wearing what is pleasing, marshalling the noble (the foe?)'. The argument for relating *ṛñjate* here to the participle *ṛñjánt-* is based on the apparent parallelism (noted by Grassmann 1873: 280 and Geldner 1951-7: 2.54 n. 5*ab*, cf. 1.124 n. 7*b*) with 1.95.7 *ubhé sícau yatate bhīmá ṛñján* 'He (Agni) terrifying, going directly, marshalls both flanks (of the army)'. But if the two passages are connected, *ṛñjate* does not provide sound evidence for finite nasal present forms from the root 'to go directly, to stretch': the expression *ṛñjate ... yatám* could be an artificial creation derived from *yatate ... ṛñján* by transference of inflections (1.95.7 does not necessarily represent the younger of the two passages, as the Triṣṭubh cadence *bhīmá ṛñján* is also found in one of the Family Books, 4.48.8).

[25] The absence of such a preverb/preposition in the contexts with nasal present middle forms such as athematic 3rd pl. *ṛñjate* makes it even less likely that the accusatives here represent an accusative of goal, and that the same inherited root is involved as in *abhí ... ṛjyate*.

[26] This is the Type 4 pattern distinguished by Jamison (1983: 147) for early Vedic involving a medially inflected *-ya-* intransitive beside an *-aya-* present with transitive value. Jamison (1983: 152) regards it as a pattern which developed mostly within Indic, but her data does not include cases such as RV *ṛjyate*/YAv. *rāzaya-*, where the *-aya-* present may have survived only in Iranian, or wholly Iranian pairs such as YAv. *draža- < *dṛnjya-/drənjaya-* (Kellens 1984: 141 n. 8).

rāzaya- and Vedic *ṛjya-* represent inherited stems,[27] we would not expect a nasal present to have existed in prehistoric Indo-Iranian, as nasal presents and *-áya-* presents were complementary in distribution.

A number of roots which form an intransitive present of similar structure to *ṛjyate* (i.e. a 'Class IV' *-ya-* present with zero-grade root and middle inflection) also possess a thematic nasal present: RV *siñcanti* (9.71.4), *siñcat* (4.43.6, etc.), etc., *sicyate* (4.49.2, 5.51.4, 9.39.5, etc.); *muñcata* (2nd pl., 3.33.13, 3.53.1, 4.12.6), etc., *múcyase* (RV 1.31.4), etc.; *vindáti* (RV 9.67.21), *ávindat* (3.34.4, 5.14.4, 5.29.3, etc.), *vidyáte* (5.44.9). As these examples show, this pattern is well-established in the RV Family Books even though from a diachronic point of view it may replace a still older pattern involving athematic nasal presents (*yunájmi*/*yujyáte*, *riṇákti*/*ricyate*). But for the roots *sic-* 'to pour', *muc-* 'to free', *vid-* 'to find' only the thematic nasal present is ever attested, and this present is normally transitive in contrast to the *-ya-* present with middle endings which is intransitive (or passive in the case of *vid-*). In the case of the root *sic-*, whose *-ya-* present is best attested, the finite forms based on the nasal stem *siñcá-* regularly have an object in the accusative (e.g. 1.85.11 *ásiñcann útsaṃ gótamāya tṛṣṇáje*, 5.83.8 *mahā́ntaṃ kóśam úd acā ní ṣiñca*). On the other hand, the active participle based on this same stem is often employed absolutely or even intransitively, an anomaly that was noted already by Grassmann (1873: 1515, *sic-* sense 6): 'sich ergiessen aus oder durch (I.), nur im Particip'. E.g. RV 5.85.6 (cf. 1.121.6, 8.49.6, 10.21.3, 10.102.11[28]):

> *imā́m u nú kavítamasya māyā́m*   *mahī́ṃ devásya nákir ā́ dadharṣa*
> *ékaṃ yád udnā́ ná pṛṇanty énīr*   *rā́siñcántīr avánayaḥ samudrám*

'Auch dieses grosse Kunststück des weisesten Gottes wagt keiner anzutasten, dass die glitzernden Ströme, *wenn sie sich ergiessen*, das eine Meer mit ihrem Wasser nicht füllen.' (Geldner 1951-7: 2.88)

In other words, from a synchronic point of view in the RV *siñcánt-* functions as a participle for the intransitive finite present stem *sicyáte* (*sicyámāna-* is

---

[27] This is likely, but cannot be proved, because no *-ya-* present corresponding to Vedic *ṛjya-* occurs in Old Iranian. It may have been replaced by medially inflected *rāzaya-* (with preverb *ham-*) which occurs in the sense 'to stretch out' in YAv. Such a double value cannot have been inherited for the stem *rāzaya-* and its intransitive use must represent an innovation. (I am grateful to Almut Hintze for drawing the value of YAv. *ham.rāzayata* to my attention.)

[28] Geldner's RV translation does not accept so many examples of this usage as Grassmann did, but Grassmann's interpretation provides the simplest solution to the problems presented by these passages.

attested, but only with the preverb *pari-*, and always constituting the second half of a Triṣṭubh) as well as for the transitive present *siñcati*. In view of this pattern of employment, it may be suggested that *sincánt-* beside *sicyáte* supplied a formal model for the creation of a 'nasal present' participle *r̥ñjánt-* beside *r̥jyate*.[29] Hence it is unnecessary to interpret this limited group of participle forms as the remnants of an obsolete nasal present from the inherited root *H₃erĝ-*: they may be explained as an Old Indic innovation,[30] and possibly an artificial poetic creation of the RV poets.[31]

---

[29] Just as in one passage *siñcánt-* has an object in the accusative (RV 5.83.6), so arguably does *r̥ñján* at RV 3.31.1, but this verse, which deals with the incest of Agni (or Dyaus), presents many difficulties and perhaps deliberate obscurities of language: *śā́sad váhnir duhitúr naptyàṃ gād | vidvā́m̐ r̥tásya dī́dhitiṃ saparyán | pitā́ yátra duhitúḥ sékam r̥ñján | sáṃ śagmyèna mánasā dadhanvé* 'Der (Opfer)lenker kam die Enkelin der Tochter zurechtweisend, (des Gesetzes) kundig, die Erkenntnis des Gesetzes hochhaltend, wobei der Vater rechtschaffenen Sinnes einverstanden war, indem er den Erguss (die Besamung) der Tochter betrieb' (Geldner 1951-7: 1.366). Translated by Renou (1955-69: 17.70): '(Agni) guide (du sacrifice) se rendit vers sa petite-fille, lui qui sait ... au moment où, effectuant le versement du (sperme dans la) fille, le Père (Ciel) s'élançait avec un esprit puissant.' However, we have the same *pāda*-final use of the active participle form *r̥ñján* agreeing with the subject of a finite verb (here *sáṃ ... dadhanve*, whose meaning and syntax is unfortunately uncertain) as we find in clear intransitive usage in three other passages (1.95.7, 4.38.7, 4.38.8). Possibly *sékam r̥ñján* is a substitute for the more 'obvious' participle with cognate accusative, *sékaṃ siñcán*?

[30] The existence of the denominative 'middle participle' *r̥ñjasāná-* may have supported the creation of *r̥ñjánt-* as the meaning and usage of the two is very similar (cf. the translation of *r̥ñjasāná-* 'going straightways' in Insler 1968: 12-16), except that they occupy different metrical 'slots'. A proportional analogy involving another *-asāná-* participle could also have produced *r̥ñjánt-*: *sáhasā : sahasāná-* : *sáhant-* = **r̥ñjasā* (> *áñjasā*) : *r̥ñjasāná-* : X => X = **r̥ñjant-* (with accent shifted in accordance with the accentuation of other thematic nasal presents to give *r̥ñjánt-*). *sáhant-* is used as an adjective (RV 5.87.5, 7.56.5, 8.40.1, 8.46.20) and does not represent a regular participle, as *sah-* normally shows middle inflection in its present forms. However, caution is necessary as it is difficult to establish whether the creation of *r̥ñjasāná-* preceded the creation of *r̥ñjánt-* or vice versa, as the analogy could have operated in the opposite direction.

[31] This proposal accords with the fact that such forms are unknown outside of the RV, but it should be regarded as tentative since our argument leaves unsolved one problem concerning the Old Indic reflexes of the inherited root *H₃erĝ-*: the origin of the *s*-stem noun **r̥ñjas-*, which has frequently been reconstructed (cf. e.g. Mayrhofer 1986: 1.55) to explain the adverbs *áñjas*, *áñjasā*, and the 'middle participle' *r̥ñjasāná-*. If there was no old well-established nasal present in Old Indic, how did the nasalized form of the root that appears in this *s*-stem arise? The whole problem could be projected back into the parent language and it might be hypothesized that IE **H₃erĝ-* had both a form with and without nasal infix (as, e.g., **le(n)gʷh-*), but this reconstruction would be entirely *ad hoc*, as no trace of a nasalized form of the root occurs elsewhere in IE. It is more likely that an *s*-stem with nasal was created within Old Indic (or Indo-Iranian) on the analogy of other *s*-stems such as RV *ráṃhas-*, 'speed, swiftness', *dáṃsas-* 'skill', *áṃhas-* 'narrowness', *jáṃhas-* 'means of movement, wing' (6.12.2), *áñjas-(pā́-)* 'ghee(-drinking)', etc. The model could have been supplied

Although the root *²ṛj-* 'to move mentally, to stimulate', which we have attempted to identify in RV verbal forms, appears to be without cognates outside Old Indic, a nominal compound containing the same root may perhaps occur in the dual form *mánaṛṅgā* (RV 10.106.8 *mánaṛṅgā mananíā ná jágmī*). Grassmann translated this epithet of the Aśvins (following Sāyaṇa) 'den Sinn oder Geist lenkend' (cf. Wackernagel-Debrunner 1896-1957: II/2.82 'den Sinn lenkend', Jamison 1983: 80 n. 5 'directing, arranging the mind'). Such a value appears to be supported by the following epithet *mananíā* which probably contains *mána-* 'mind' (= *mánas-*) + *nī́-*, a verbal noun from the root *nī́-* 'to lead', but the whole passage is obscure, involving a word-play on the sounds *-ph-* and *-r-*. *-ṛṅga-* was explained by Wackernagel-Debrunner (1896-1957: I.161-2 and II/2.82) as a verbal noun from *ṛj-/ṛñj-*, which shows a velar stop, in place of the palatal they expected, as a result of analogy with roots ending in a voiced labiovelar or velar stop where an alternation *g/j* was regular. In other words, if *-ṛṅga-* is connected with the root *ṛj-/ṛñj-* of the *ṛjú-* family, it has to be consiered to contain unetymological *-g-*, just as, e.g., RV *sárga-* 'emission, flood' from *sṛj-* 'to pour' (cf. Avestan *harəz-*) or Classical Skt. *yāga-* 'offering' from *yaj-* 'to worship' (cf. Av. *yaz-*, OP *yad-*, Gk. ἅζομαι). However, *-ṛṅga-* must be a verbal noun derived from the nasal present stem,[32] and as such would represent the only case of this type where an analogical replacement of *-j-* by *-g-* has to be postulated.[33] Moreover, it is hard to understand from where the guttural could have been extended, if it did not originally occur in some forms belonging to this nasal present. A connection with the root *²ṛj-* which is under discussion would also be

---

particularly by those roots where from a synchronic point of view the *s*-stem noun shows a nasal but the corresponding 'Caland type' adjective does not (*rámhas-/raghú-*, *dáṃsas-/dasrá-*, *dasmá-*); and the motivation for its creation could have been to distinguish a noun **ráñjas, ṛñjásas* 'straightness, directness' from the original *s*-stem *rájas-* which had come to be used in the sense 'space' (cf. Burrow 1947-8: 648, Mayrhofer 1986-96: 2.426) on the connection between the latter noun and **H₃erǵ-*). It may also be observed that the root *raj-* 'to redden' becomes *rañj-* from Epic Skt. onwards, and in **ṛ́ñjas-* we may have an earlier case of a development which took place in roots of this phonological shape (Jamison 1983: 152 attributes the later nasalization of *raj-* to the analogy of other roots ending in palatal affricates, most of which have a nasal).

[32] This type of Vedic verbal noun is also represented by *agnim-indhá-* 'kindling the fire' (RV 1.162.5), *viśvam-invá-* 'driving all' (RV Family Books, I and X), *ā-daghná-* 'reaching the mouth' (RV 10.71.7), AV *bhūmi-dṛmhá-* 'stabilizing the ground', etc.; cf. YAv. *yimō.kərənta-* 'cutting-Yima (in two)' (Wackernagel-Debrunner 1896-1957: II/1.181-2).

[33] In addition if *mánaṛṅga-* is indeed a determinative compound (with shifted accent because it is a vocative in *pāda*-initial position), the traditionally assumed replacement of *-j-* by *-g-* would run counter to the normal Indic development whereby a palatal replaces a guttural in thematic verbal nouns with *nomina agentis* value, rather than vice versa (cf. Wackernagel-Debrunner 1896-1957: I.149-50).

viable from a semantic point of view: a meaning 'stimulating the mind, moving the mind' would be an appropriate epithet for the Aśvins and would also fit with the following *mananíā*. Thus if this account of the second member *-ṛṅga-* is accepted, *mánaṛṅga-* may provide a fragment of phonological evidence to support the hypothesis that the root *²ṛj-* had an allomorph *ṛg-* (i.e. that this root, in contrast to the root that continues IE *$*H_3erĝ$-*, ended in a plain velar or labiovelar).

At this stage we may return to the discussion of *ṛgmín-* at the point where it had been argued that, because of the existence of RV *ṛgmíya-* and the difficulty of deriving *ṛgmín-* from *ṛk-* + *-mín*, it is necessary to reconstruct a primary noun **ṛgma-* as the source of both these secondary derivatives. **ṛgma-*, it was argued, could only derive from a root with a voiced final stop **ṛj-/ṛg-*. That this root is *²ṛj-*, which we have identified in the verbal forms and more tentatively in the compound *mánaṛṅga-*, is suggested by the following set of parallel collocations:

  girā́ ... ṛgmíṇaḥ (RV 9.86.46)
  ṛñjate girā́ (RV 2.2.5)
  ṛñjase girā́ (RV 4.8.1, 6.15.1)
  gīrbhíḥ ... ṛgmíyam (RV 6.45.7, 1.9.9, 1.51.1)

  suvṛktíbhiḥ ... ṛgmíyam (RV 1.62.1, 8.40.10)
  suvṛktíbhiḥ ... ṛñjase (RV 6.15.4)

In the case of both the verbal and nominal forms there is a root of the shape **ṛj-/ṛg-* which is associated with precisely the same nouns in the instr. meaning 'song, fine verse'. In other words, it is possible to identify an expression 'to move, to stimulate by means of song/fine verses', which was part of RV poetic diction and which was expressed by means of a derivative from the root *²ṛj-* + instrumental of a noun *gír-* or *suvṛktí-*.

The contexts for *ṛgmíya-* are clearer and these will be discussed first. If the primary derivative **ṛgma-* was a nomen actionis from the root *²ṛj-*, we would expect it to have meant 'stimulation, action of moving (the mind, emotions, etc.)'. On the morphological and semantic parallel afforded by *stóma-* 'praise' : *stómia-* 'to be praised', or inherited *yajñá-* 'worship' : *yajníya-* 'to be worshipped' (cf. Av. *yasna-/yesniia-*) an original sense 'to be moved, to be stimulated' may be proposed for *ṛgmíya-*. Such a meaning is possible in all 11 passages where the word occurs, e.g. RV 6.45.7:

brahmā́nam bráhmavāhasaṃ   *gīrbhíḥ sákhāyam r̥gmíyam*
gā́ṃ ná doháse huve

'I call on the Brahmin (Indra) who receives respect from the sacred formula, the friend who is to be stimulated by songs, like a cow to give milk.'

The syntagma with *girā́*, *gīrbhíḥ*, *suvr̥ktíbhiḥ* that is found in 5 out of the 11 passages where *r̥gmíya-* occurs is sometimes discontinuous, but it is always recognizable, e.g. RV 8.40.10:

tā́ṃ śiśīta suvr̥ktíbhis   *tveṣáṃ sátvānam r̥gmíyam*

'Sharpen him (Indra) with fine verses, the terrifying warrior (who is) to be stimulated.'

Note here the clear context involving stimulation (expressed by *śiśīta*).[34]

   Just as verb forms from *²rj-* are associated with forms from *stu-* 'to praise', so is *r̥gmíya-*. Compare RV 8.4.17 *vémi tvā pūṣan r̥ñjáse | vémi stótava āghr̥ṇe* 'I seek you, O Pūṣan, for stimulation, I seek you, O glowing one (?), for praising' with 8.39.1 *agním astoṣy r̥gmíyam* 'I have praised Agni (who is) to be stimulated'.

   Another opening verse combines *r̥gmíya-*, *suvr̥ktíbhiḥ*, and a form from *stu-* all in the same *pāda* (RV 1.62.1):

prá manmahe śavasānā́ya śū́ṣam   *aṅgūṣám gírvaṇase aṅgirasvát*
suvr̥ktíbhi stuvatá r̥gmiyā́ya   *árcāma arkáṃ náre víśrutāya*

Geldner (1951-7: 1.80) translates 'Wir ersinnen für den Mächtigen eine Stärkung, für den Lobbegehrenden ein Loblied in der Weise der Angiras'. Auf ihn, der für den Sänger in schönen Preisreden zu preisen ist, wollen wir einen Lobgesang singen, auf den berühmten Herrn', but 'to be moved with fine verses

---

[34] Likewise RV 1.51.1 *abhí tyám meṣám puruhūtám r̥gmíyam | índraṃ gīrbhír madatā vásvo arṇavám* 'Den bekannten Widder, den vielgerufenen, preiswürdigen Indra ergötzet mit Lobreden, der ein Strom des Guten ist' (Geldner 1951-7: 1.61-2); here also *r̥gmíyam* may be translated as 'to be stimulated' but *gīrbhír* must depend syntactically on the finite verb *madatā*. However, at 1.9.9 *vásor índraṃ vásupatiṃ | gīrbhír gr̥ṇánta r̥gmíyam | hóma gántāram ūtáye* 'den Herrn des Guts, den preiswürdigen Indra mit Reden lobend, der auf den Ruf Kommt, um Gunst zu erweisen' (Geldner 1951-7: 1.10), it would be possible to construe *gīrbhír* with both *gr̥ṇánta* and *r̥gmíyam* 'greeting with songs, to be stimulated (with songs)'. Outside the Family Books the collocation appears to be preserved in transposed form, possibly because *r̥gmíya-* was acquiring a new meaning (see below).

for the praiser' is possible for *pāda* c.[35] However, here *suvṛktíbhis ... ṛgmiyā́ya* may indeed have been intended as a sort of *figura etymologica* parallel to *arcāma arkám* of *pāda* d, even though in origin it was not such.[36]

As the noun **ṛgma-* had already been lost at a prehistoric date and the verb *²ṛj-* became obsolete in the post-RV language (it does not occur at all in the AV and YV), a secondary association between *ṛgmíya-* and the root *ṛc-* may have already been developing in the language of the later RV poets.[37] The connection would have been possible because a stem-form *ṛg-* occurred in e.g. *ṛgbhíḥ* (RV 2.35.12), instr. pl. of the root noun *ṛ́c-*. It is noteworthy that *ṛgmíya-* is employed in initial verses in RV Maṇḍalas I and VIII by the same poets who use other words in their opening invocations that seem to recall *ṛc-*, e.g. Indra's epithet *ṛ́cīṣama-*, whose meaning and etymology is quite unclear. It is perhaps not too daring to speculate that *ṛgmíya-* was used in such contexts in a punning fashion.

By later Vedic the association of *ṛgmíya-* with *ṛc-* is the only one, and both its accentuation and value have changed: TS 6.6.2.1 *ṛ́gmiya-* and MS 4.8.4 *ṛ́gmya-* mean 'consisting of a ṚgVeda verse, of Ṛc-verses'. As pointed out by Hoffmann (1960: 13-14) the same word occurs at JB 2.73 in the form *+ṛkmya-*. Here, it may now be suggested, we have the final stage in the reinterpretation of this word as a derivative from *ṛc-*, and this has resulted in a further modification to its phonology.[38]

Turning at last to the passages containing *ṛgmín-*, if it is correct that we have an *-ín* derivative from **ṛgma-* 'stimulation, action of moving', we should expect this word to have meant 'possessing stimulation, stimulating' just as from *śúṣma-* 'vehemence' we find *śuṣmín-* 'possessing vehemence, vehement'. Hence we may now translate RV 9.86.46bc:

---

[35] Cf. RV 8.23.2-3 *dāmā́naṃ viśvacarṣane | agníṃ viśvamano girā́ | utá stuṣe víspardhaso ráthānām | yéṣām ābādhā́ ṛgmíya | iṣáḥ pṛkṣáś ca nigrábhe* 'Den Geber Agni mit Lobrede, du allbekannter Viśvamanas, und preise die wetteifernden (Rosse) der wagen, Deren preiswürdiger Ungestüm die Speisen und Stärkungen erraffen soll!' (Geldner 1951-7: 2.328-9). Here we find *girā́* (possibly elliptical for *ṛñjase girā́*), *stuṣe*, and *ṛgmíya-* in successive clauses: the sense proposed here for *ṛgmíya-* would enable us to translate the relative clause as 'of which the impetuosity is to be stimulated for holding fast strengths and refreshments'.

[36] If the explanation offered by Thieme (1951: 172-3) of *śūṣam < *(p)śu-san-* is correct, we have a parallel case in *śavasānā́ya śūṣám* in *pāda* a, where a false figura etymologica is employed. But some scholars believe these two words do contain the same root (cf. Mayrhofer 1986-96: 2.652).

[37] The poet of 1.62 has innovated with the form and accent of *ṛgmiya-*, since elsewhere in the RV only the forms *ṛgmíyam* (x 8) and *ṛgmíyaḥ* (x 2) occur (always at the end of a *pāda*).

[38] Bailey (1958: 152) draws attention to a parallel (reported by S. P. Pandit) in the case of the RV adjective *śagmá-* (discussed above). Present day Ṛgvedins regularly recite *śakma-*, because of Sāyaṇa's gloss *śakta-*, which created a new etymology by connecting this word with the root *śak-* 'to be able' after the root *śag-* had disappeared in Indic.

aṃśúṃ rihanti matáyaḥ pánipnataṃ   girā́ yádi nirṇíjam ṛgmíṇo yayúḥ

'The poems lick the wonderful plant when those who stimulate through song have gone to the garment (of Soma).'

Here *girā́ ... ṛgmíṇo yayúḥ* must express a similar idea to *ṛñjate girā́* (RV 2.2.5). The *ṛgmíṇaḥ* are the composers of the poems (*matáyaḥ*) as Geldner and Renou believed (see the beginning of this paper for their translations), but they are described as '(those) stimulating through song' because their poetry stimulates Soma just as mother cows (in other words, the frequent RV equation *dhenú- : matí-*) stimulate their calves by licking.[39]

RV 1.100.4

só áṅgirobhir áṅgirastamo bhūd   vṛ́ṣā vṛ́ṣabhiḥ sákhibhiḥ sákhā sán
ṛgmíbhir ṛgmī́ gātúbhir jyéṣṭho   marútvān no bhavatu índra ūtī́

'He is the best Angiras among the Angirases, the bull among bulls, the friend among friends, the stimulating one among those who stimulate, the best among ways - may Indra accompanied by the Maruts be with help for us.'

Here the juxtaposed nominative and instrumentals form a progression from Indra's Angiras (= Bṛhaspati) role[40] to one where he represents strength (as a bull), to that of a friend, to that of one who possesses the power to stimulate and so is the best of ways. Each step leads to the request in the refrain line for Indra in association with the Maruts to help. There is more point in mentioning his ability to stimulate at the penultimate stage of the progression than his ability to sing verses. As *ṛgmíbhir* refers to the Maruts it should be compared not with their epithet *ṛ́kvabhiḥ* but with *iṣmínaḥ* in three other hymns (1.87.6, 5.87.5, 7.56.11). Here we have two morphologically parallel derivatives from two roots that mean 'to move': *²rj-/ṛg-* 'to move mentally, to stimulate', and *iṣ-* 'to impel, to send, to move physically'. A primary noun from which *iṣmín-* derives may be

---

[39] The converse of this idea, Soma's need for stimulation, is expressed by *ṛgmíya-* in another Jagatī hymn of RV IX: cf. 9.68.6 *(manīṣíṇaḥ) ... tám marjayanta suvṛ́dhaṃ nadī́ṣu ā́m̐ | uśántam aṃśúm pariyántam ṛgmíyam* '(The poets) wash him of good growth (Soma) in the rivers, the willing plant (who) moves around (and) is to be stimulated'.

[40] This role, of course, involves singing or reciting: cf. Schmidt (1968: 37-8). Therefore on this ground alone *ṛgmíbhir ṛgmī́* might be expected to refer to a different quality of Indra's.

lacking because it may have been created on the analogy of *r̥gmín-* directly from the verb *iṣ-, iṣṇā́ti*.[41]

All these facts point to an early loss of the original meaning of *r̥gmín-*, *r̥gmíya-*. As *²r̥j-* became obsolete even by the later Vedic language and its only continuation is in Middle Indo-Aryan, not Classical Sanskrit, it is hardly surprising that the original meaning was lost to the Indian tradition which sought to link its derivatives with more familiar roots. The conclusions reached in this paper about the original sense and employment of *r̥gmín-*, *r̥gmíya-* on the one hand, and *r̥ñjate* and related verb forms on the other mutually support each other, as they make it possible to identify a root *²r̥j-* and establish its meaning via both nominal and verbal forms.[42]

References

Bailey, Harold W. (1958). 'Iranica et Vedica: 1. Iranian *sagra-*, Old Indian *śagmá-*'. *IIJ*, 2: 149-52.

Bartholomae, Christian (1904). *Altiranisches Wörterbuch*. Berlin.

Benveniste, Émile (1969). *Le vocabulaire des institutions indo-européennes* (2 vols.). Paris.

Bloomfield, Maurice (1910-11). 'Some Rig-Veda repetitions: 1. The meaning and etymology of *iṣmín*'. *JAOS*, 31: 49-51.

Böhtlingk, Otto, and Roth, Rudolph von (1855-75). *Sanskrit Wörterbuch*. St. Petersburg.

Burrow, Thomas (1947-8). 'Sanskrit *rájas*'. *BSOAS*, 12: 645-51.

Cowgill, Warren (1968). 'The aorists and perfects of Old Persian'. *KZ*, 82: 259-68.

Geldner, Karl F. (1951-7). *Der Rig-Veda aus dem Sanskrit ins Deutsche übersetzt* (Harvard Oriental Series, 33-5). Cambridge, Mass.

——, and Pischel, Richard (1901). *Vedische Studien III*. Stuttgart.

Gotō, Toshifumi (1987). *Die 'I. Präsensklasse' im Vedischen*. Wien.

Grassmann, Hermann (1873). *Wörterbuch zum Rig-Veda*. Leipzig.

---

[41] Mayrhofer (1986-96: 1.202) suggests the analogy of *bhāmín- : bhā́ti* 'shines', but *r̥gmín- : *r̥ṅkte, r̥ñjate* would have provided a closer morphological and semantic model. It may also now be proposed that the late Vedic *vāgmín-* (Ś.Br.+) was derived from *vā́c-* 'speech' on the analogy of *r̥gmín-* at a stage when the latter had been reanalysed and reinterpreted in its *mantra* occurrences as a *-mín-* derivative from the root noun *r̥c-*.

[42] It is outside the scope of this paper, but it would be worth investigating whether *r̥jīṣá-*, *r̥jīṣín-*, the epithet of Indra which is connected with his Soma drinking, might also be a derivative of the same root *²r̥j-*.

Haebler, Claus (1968). 'Pā. *iñjati*, buddh. h. Skt. *iñjate* : Ved. *ṛñjate*, eine mittelindisch-vedische Isolexe', in J. C. Heesterman, G. H. Schokker, and V. I. Subramoniam (eds.), *Pratidānam* (Festschrift F. B. J. Kuiper). Den Haag/Paris, 283-98.

Haudry, Jean (1977). *L'emploi des cas en védique*. Paris.

Hoffmann, Karl (1960). 'Textkritisches zum Jaiminīya-Brāhmaṇa'. *IIJ*, 4: 1-36 (repr. in Johanna Narten (ed.), *Aufsätze zur Indoiranistik I*. Wiesbaden, 1975, 77-112).

Insler, Stanley (1968). 'Vedic *áñjasā*, *ṛñjasāná-* and the type *sahasāná-*'. *KZ*, 82: 1-23.

— (1999). 'Vedic *dṛṣád-*', in H. Eichner and H. C. Luschützky (eds.), *Compositiones Indogermanicae in memoriam Jochem Schindler*. Praha, 163-4.

Jamison, Stephanie (1983). *Function and Form in the -áya-Formations of the Rig Veda and Atharva Veda*. Göttingen.

Joachim, Ulrike (1978). *Mehrfachpräsentien im ṚgVeda*. Frankfurt a. M./Bern/ Las Vegas.

Kellens, Jean (1984). *Le verbe avestique*. Wiesbaden.

— (1995). *Liste du verbe avestique*. Wiesbaden.

Lamberterie, Charles de (1990). *Les adjectifs grecs en -ύς* (2 vols.). Louvain-la-Neuve.

Lubotsky, Alexander (1981). 'Grk. *pēgnumi* : Skt. *pajrá-* and the loss of laryngeals before Mediae in Indo-Iranian'. *MSS*, 40: 133-8.

— (1997). *ṚgVedic Word Concordance*. New Haven.

Macdonell, Arthur Anthony (1910). *Vedic Grammar*. Strassburg.

— (1916). *Vedic Grammar for Students*. Oxford.

Mayrhofer, Manfred (1986-96). *Etymologisches Wörterbuch des Altindoarischen* (2 vols.). Heidelberg.

Narten, Johanna (1980). 'Ved. *āmáyati*, *āmayāvin-*'. *StII*, 5/6 (Festschrift P. Thieme): 153-66.

Neisser, Walter (1924). *Zum Wörterbuch des ṚgVeda* (2 vols.). Leipzig.

Peters, Martin (1986). 'Probleme mit anlautenden Laryngalen'. *Sprache*, 32: 365-83.

Renou, Louis (1955-69). *Études védiques et pāṇinéennes* (17 vols.). Paris.

Rix, Helmut (1998). *Lexikon der indogermanischen Verben*. Wiesbaden.

Santucci, James A. (1982). 'The relationship of √Ṛj (√Arj) and √Sṛj in the Veda', in T. N. Dharmadhikari (ed.), *Golden Jubilee Volume*. Poona, 241-54.

Scharfe, Harmut (1985). 'The Vedic word for "king"'. *JAOS*, 105: 543-8.

Schlerath, Bernfried (1984). 'Beobachtungen zum Wortfeld "singen, preisen, rufen, verkünden" im RV'. *MSS*, 44: 191-214.

Schmidt, Hanns-Peter (1968). *Indra und Bṛhaspati*. Wiesbaden.

Schrijver, Peter (1991). *The Reflexes of the Proto-Indo-European Laryngeals in Latin.* Amsterdam/Atlanta.

Sihler, Andrew (1977). 'The etymology of PIE *rēǵ-* "king", etc.'. *JIES*, 5: 221-46.

Strunk, Klaus (1987). 'Further evidence for diachronic selection: Ved. *rāṣṭi*, Lat. *regit*, etc.', in G. Cardona and N. H. Zide (eds.), *Festschrift for Henry Hoenigswald*. Tübingen, 385-92.

Thieme, Paul (1951). 'Etymologische Vexierbilder'. *KZ*, 69: 172-8 (repr. in *Kleine Schriften I*. Wiesbaden, 1971, 57-63).

Turner, Ralph L. (1966). *A Comparative Dictionary of the Indo-Aryan Languages*. Oxford.

Wackernagel, Jacob, and Debrunner, Albert (1896-1957). *Altindische Grammatik. I: Lautlehre, II/1: Nominalkomposition, II/2: Die Nominalsuffixe.* Göttingen.

Werba, Chlodwig H. (1992). 'Prakritic wordforms in the ṚgVedaSaṃhitā'. *WZKSO*, 36: 11-18.

— (1997). *Verba Indoarica I: Radices Primariae*. Wien.

Whitney, William Dwight (1889). *Sanskrit Grammar* (2nd edn.). Cambridge, Mass.

—, and Lanman, Charles R. (1905). *Atharva-Veda Saṃhitā Translated with a Critical and Exegetical Commentary* (Harvard Oriental Series, 7-8). Cambridge, Mass.

Periphrastic causatives with ποιέω in Ancient Greek prose

Diana Gibson

1. Given the recent interest in causative verbs by generative and typological linguists, there is a surprising lack of research on causative verb formation in Ancient Greek. While it is frequently acknowledged that Greek inherited a number of morphological causative formations from Indo-European (Schwyzer-Debrunner 1939-71: 2.222; Rix 1992: 195), the fact that Ancient Greek developed its own, very productive method of creating causatives periphrastically with the verb ποιέω has been neglected.

One of the earliest scholars to study causative verbs in Ancient Greek, W. Kühne in his 1882 *Das Causativum in der Griechischen Sprache*, simply dismissed the periphrastic type of causatives, suggesting they were of little linguistic interest: '[…] die zusammensetzungen und umschreibungen mit ποιεῖν, wie sie namentlich in der späteren sprache immer häufiger werden, da diese reinen umschreibungen des gedankens kein sprachliches interesse mehr bieten' (Kühne 1882: 2). The standard grammars of Classical Greek make almost no explicit reference to the origins and development of the causative function of ποιέω.[1] Although Wackernagel remarks on the particular syntax of the accusative + infinitive construction, he is rather elusive about the status of the construction in the Classical period, suggesting only that the construction is common in later Greek:

> Wenn z.B. bei Verben, die irgend ein Veranlassen oder Bitten ausdrücken, der Veranlasste im Akkusativ gegeben wird und dann im Infinitiv die Tätigkeit genannt ist, zu der er

---

[1] Schwyzer-Debrunner (1939–71: 2.220) concur that the construction is common in late Greek; Meier-Brügger (1992) mentions the construction not at all, while listing the extensive literature on the infinitive. Humbert (1960: 258-9) does mention causativity briefly, but only in the context of a note on the passive. Meillet (1965), Chantraine (1991), Rix (1992), and Cooper (1998) are unhelpful; Kühner-Gerth (1966: 2.2.10ff.) is unsurprisingly the most informative: 'Infinitiv ohne Artikel als Ergänzung von Satzgliedern: die Verben des Machens, Bewirkens, als: ποιῶ, πράττω, διαπράττομαι, κατεργάζομαι, κατασκευάζω, καθίστημι' etc.; further, Kühner-Gerth (1966: 2.1.318, §411) under 'Doppelter Akkusativ': 'Erstens steht ein Akkusativ des Objekts und ein Akkusativ des Prädikats bei den §355 angeführten Verben, wenn sie aus Passiven in Aktive verwandelt werden, also bei den Ausdrücken: a) ich mache einen oder etwas zu etwas, als: ποιῶ, ποιοῦμαι (*mihi facio*), τίθημι, *reddo*, καθίστημι…'; also Kühner-Gerth (1966: 2.1.44, §356 A.2; 2.2.372-3, §552).

veranlasst wird, so haben wir gewissermassen das Subjekt der Infinitiv im Akkusativ vor uns; z.B. in dem homerischen Satze ψ 258 σὲ θεοὶ ποίησαν ἱκέσθαι οἶκον "die Götter machten dich nach Hause zurückkehren" ist σέ zugleich Objekt von ποίησαν und Subjekt von ἱκέσθαι. Vieles Derartige findet sich noch im spätesten Griechisch.

<div align="right">(Wackernagel, 1926-8: 1.263)</div>

Despite Kühne's suggestion that causatives with ποιέω plus infinitive are a phenomenon of later Greek, some of the oldest Greek literature contains periphrastic causatives — including the passage from Book 23 of Homer's *Odyssey* discussed above by Wackernagel, in which Penelope acknowledges how the gods caused Odysseus to come home.

(1)   σὲ θεοὶ ποίησαν ἱκέσθαι οἶκον (*Od.* 23.258)
        'the gods made you come home'

This paper will trace out the distribution and development of the ποιέω constructions in a variety of Classical Greek prose texts, showing that, in contrast with the current assumptions, the constructions flourished sometime between the 5th and 4th centuries B.C.E. and became a feature of the philosophical language of Plato and Aristotle. The constructions with ποιέω are also present in verse, though they are less frequent there. The distribution of the verse data suggests that diffusion of the ποιέω construction began in colloquial language, a hypothesis discussed in my forthcoming Oxford University DPhil thesis, *Causative Verbs in Ancient Greek*.

2. Periphrastic causatives in Greek and many other languages are serial verb constructions consisting of a causative main verb plus an object and a verbal complement. In typical constructions, a causer makes a causee carry out some action or enter into some state of being. The standard causative verb in Ancient Greek is ποιέω, one of the most common verbs in the language and the standard verb for 'making' and 'doing'. The use of this verb is quite unsurprising, since typological research has shown that, across languages, lexically neutral verbs meaning 'make' are most frequently used to express causativity in periphrastic constructions, with many eventually becoming grammaticalized affixes (Moreno 1993, Song 1990). Classical Greek has a number of other verbs which mean 'make' in certain contexts: πράττω, δράω, παρασκευάζω, καθίστημι, τίθημι, etc. Some of these verbs can also serve as causative main verbs in prose and verse, though in the case of prose they are far less common than ποιέω and have different, less bleached, semantics.

By far the most common type of periphrastic causative in the prose texts surveyed here is ποιέω plus the accusative and infinitive. The Homeric example above is matched by many similar examples in other authors, e.g.

(2)   ὅλον ἐποίησε φεύγειν τὸ τῶν ἐναντίων (Xen. *Hell.* 7.5.24)
        'he made the whole of the enemies flee'

There are a total of 497 occurrences in my corpus,[2] distributed as follows: Thucydides 3, Lysias 4, Herodotus 13, Xenophon 30, Demosthenes 35, Isocrates 56, Aristotle 157, and Plato 199 (see Tables 1 and 2).

Table 1: Periphrastic causatives with ποιέω by Author as % of Total Data Set
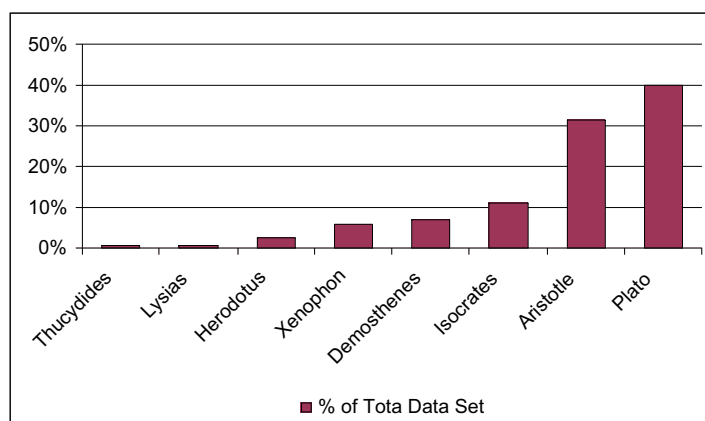

As might be expected from the relative size of their corpora, the majority (over 70%) of constructions in the data set come from Aristotle and Plato. Furthermore, Plato and Aristotle contain *proportionally* far more periphrastic causative constructions with ποιέω than most other prose authors (Table 2).

Table 2 makes clear that Thucydides has a disproportionately low number of periphrastic causatives with ποιέω, while Plato and Aristotle have an extremely high proportion — around 3.5 occurrences per 10,000 words. However, it is important to note that Isocrates has proportionally nearly as many causatives with ποιέω as Aristotle and more than Plato, suggesting that the growth of these causatives was not just a philosophical phenomenon but was becoming widespread in some forms of prose.

---

[2] For Lysias, Herodotus, Thucydides, Xenophon, Demosthenes, and Plato, all quotations follow the OCT texts. For Isocrates I have used the Teubner text by G. E. Benseler and F. Blass, for Aristotle the Bekker edition. I have included data from all works in each edition, regardless of whether they are uncontroversially ascribed to that author.

Table 2: Frequency of periphrastic causatives with ποιέω per 10,000 words by Author
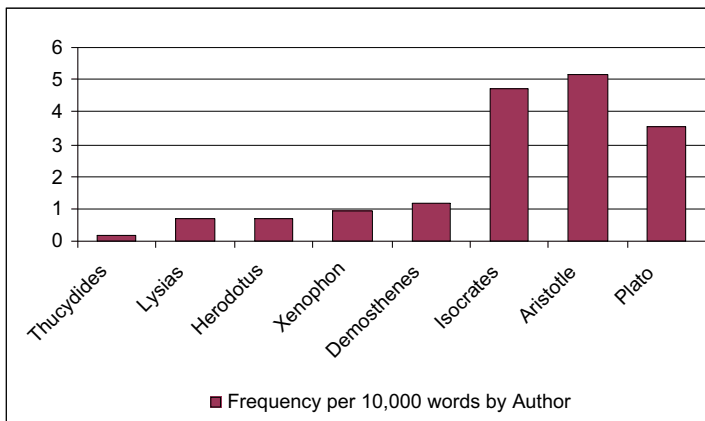
Source for word count data: Perseus Project

3. Now that we have seen the rough distribution of these constructions, we can begin to discuss their meaning and form. The examples from these authors have precisely the meanings we would expect from causatives: the causee is made by the causer to do some action, or enter into some state of being. Take for instance some causatives from the corpus of Plato:

(3)   ἀλλ᾽ εἴ μοι ἐθελήσαις αὐτὸν ποιῆσαι εἰς λόγους ἐλθεῖν (Pl. *Lys.* 206c)
       'but if you want to make him enter into conversation with me'

(4)   ἡ σωφροσύνη οὐ ποιήσει αὐτὸν γιγνώσκειν (Pl. *Chrm.* 170d)
       'but wisdom will not make him become knowledgeable'

First, the syntactic marking is what we might normally expect in Greek: the causer most often appears in the nominative case as the subject of the entire construction. The causee, which is the subject of the underlying clause, appears in the accusative case as the object of the causative verb. As is the rule in Greek, the subject of the infinitive normally appears in the accusative case when it is different from the subject of the main verb. The verbal complement of the causative verb normally appears in infinitive form, as in English and many other languages. This pattern holds true for the vast majority of the data and is unproblematic.

Second, at least two different types of causative meaning can be expressed by the same syntax. The first passage from Plato's *Lysis* suggests an element of human compulsion, or at least unwillingness on the part of the causee: 'but if you want me to make him enter into conversation' (even though he's not inclined). On the other hand, the passage from the *Charmides*, 'wisdom will not make him know' suggests a more neutral kind of natural causation, in which a

different state of being is created, but the willingness of the causee is not at issue.

3.1. These are rather subjective ways of describing the semantics of periphrastic constructions; ideally we would like to be able to define the relationship between the causer and the causee a bit more precisely. One way to do this is to consider the 'Animacy' of the causers and causees. Across the data set, approximately 42% of subjects of the causative ποιέω can be characterized as +Animate, while the rest are –Animate. However, this overall total disguises a sharp contrast between authors: the vast majority of subjects in the corpus of Aristotle are inanimate. The percentage of +Animate subjects is also very low in Plato compared to the rest of the authors surveyed, though it is still significantly higher than in Aristotle. As we might expect, Isocrates also has a relatively low proportion of +Animate subjects for a non-philosophical author; even so, +Animate subjects characterize more than 60% of the occurrences of his causative constructions (Table 3).

Table 3: Percentage of +Animate Subjects/All Subjects by Author
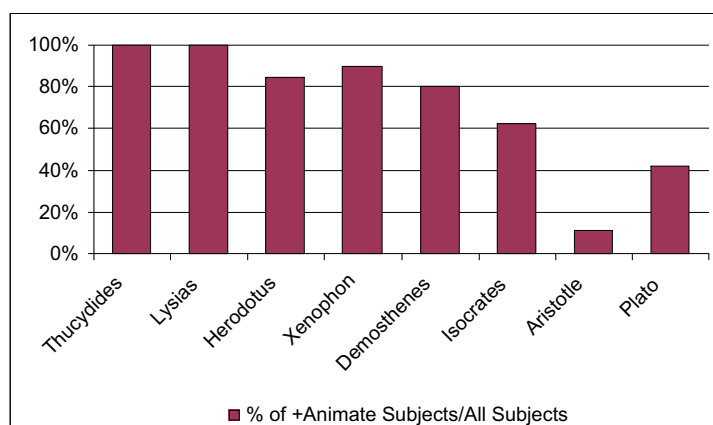
To summarize, most prose authors besides Plato and Aristotle have primarily +Animate causers. The reason for this may be simply that philosophers have a greater need to discuss abstract ideas.

   Rather obviously, since the subject of a causative construction precipitates the action or change of state described, a –Animate subject rules out many semantic possibilities including verbal order, command, and some physical violence. Accordingly in Aristotle, causative constructions with ποιέω occur mostly in the context of the physical world, where natural cause is being described. Note that in example (6) below, the interpretation of δακρύειν as a nominal is also possible. Nevertheless, I have included examples of this type in the data set and discuss them further below in the context of explicit and non-explicit causees.

(5)   ὑπὸ τοῦ πνεύματος αἱ φλέβες ἐμφυσώμεναι τοὺς πόρους συμμεμυκέναι ποιοῦσι
      (Arist. *Pr.* 868a)
      'the veins, inflated by the breath, cause pores (of the skin) to close'

(6)   καὶ γὰρ τὸ θερμὸν ποιεῖ δακρύειν, οἷον ὁ ἥλιος (Arist. *Pr.* 882a)
      'for the heat also causes (one) to weep, like the sun'

The distribution of causative subjects in Plato is also skewed to the –Animate side, though less severely than in Aristotle. And although the pragmatic context of Plato is much less easy to generalize, we can nevertheless attribute a large number of –Animate subjects to the abstract nature of the philosophical discussion, e.g.,

(7)   ἀλλὰ τὸ πρέπον, ὦ Σώκρατες, καὶ εἶναι καὶ φαίνεσθαι ποιεῖ καλὰ παρόν. (Pl. *Hp. Ma.* 294c)
      'But the appropriate thing, Socrates, makes things both be and appear beautiful by its presence.'

In Plato, as in Aristotle, a whole variety of inanimate causative agents are possible (e.g. 'medicine', 'wisdom', 'experience'). But outside of Aristotle and Plato, –Animate causative agents still exist, even if far less frequently as in Isocrates. We might wonder if the same is true of the 'causee' objects of the causative matrix verbs.

3.2. The constructions with ποιέω can be broadly divided into two groups — those with an explicit causee in the accusative case (or exceptionally the dative or the genitive), and those with no obviously specified object at all. Passages without explicit objects tend to be of the type already mentioned above in Aristotle's *Problemata* 882a8-9 'for the heat also causes (one) to weep, like the sun' (Table 4).

Table 4: Percentage of Explicit Objects/All Objects
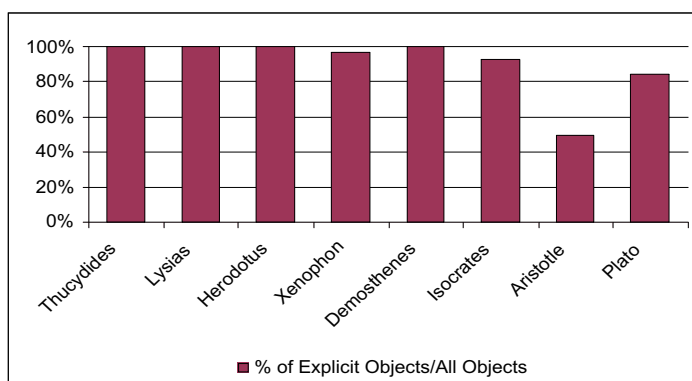
An extremely high percentage of all authors apart from Plato and Aristotle have an explicit object — Isocrates has the next lowest proportion at 93%. The number of explicitly stated causees in the Platonic corpus is slightly lower at 84%. In Aristotle, however, the picture is quite different, with an explicit object occurring in the text only 50% of the time. When the object is explicitly stated, it may be animate or inanimate. Table 5 shows the proportion of all +Animate objects per explicit object for each author.

Table 5: Percentage of +Animate Objects/Explicit Objects:
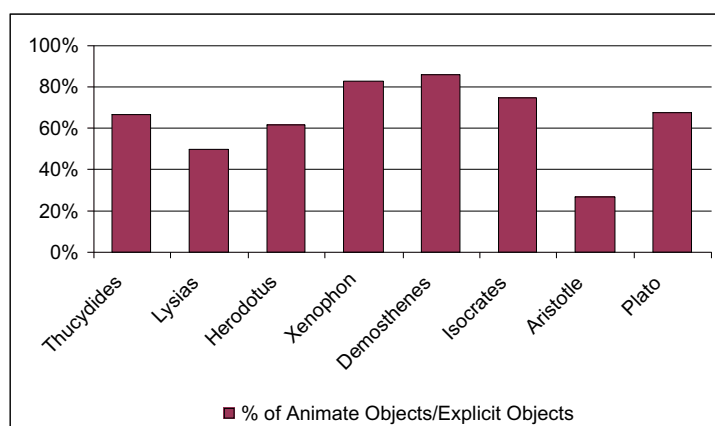
Once again, the picture is markedly different for Aristotle than for the other authors. Of the objects that are actually explicit in Aristotle, only a very small proportion (approximately 27%) are +Animate. Here a distinction between Plato and Aristotle emerges: not only does Plato have a higher proportion of explicit objects than Aristotle, but a far greater proportion of those explicit objects (approximately 67%) are +Animate than in Aristotle.

The causatives in Aristotle with non-explicit objects demand further explanation. These constructions cause problems of translation, since it is not always clear whether a passage like Arist. *Pr.* 863b ἢ ποιήσει ἀλεάζειν should mean '(it is heat) which causes (the general phenomenon of) boiling' or '(it is heat) which causes (something) to boil' with an implied object. While the latter interpretation resembles a periphrastic causative construction with an elided indefinite object, the former implies a nominal use of the infinitive with the elision of the article. In fact, passages elsewhere in Aristotle show a very similar use of the articular infinitive:

(8) τὸ γὰρ φῶς ποιεῖ τὸ ὁρᾶν. (Arist. *Sens.* 447a)
     'For light causes the seeing.'

We can conclude that Aristotle shows a somewhat specialized use of the periphrastic causative which may be related to his particular use of the articular infinitive.

3.3. Another way to understand the semantics of causative constructions is to see what type of verb is used to indicate the action or change of state which the causer makes the causee carry out. The simplest way to do this is to categorize verbal complements by their transitivity (transitive/intransitive) or by their rough semantic type — for the latter I have used the categories 'action' (*to run*, *to flee*, etc.) versus 'change of state' (*to grow old*, *to become white*, etc.).

While it is true that in some languages causatives can only be formed to intransitive underlying clauses, in most languages periphrastic causatives can be formed to base constructions with both intransitive and transitive verbs. There is no typological evidence for a language forming causatives to transitives but not to intransitives (Dixon and Aikhenvald 2000: 45). Ancient Greek allows causatives to both transitives and intransitives, though causatives to intransitives make up the majority of occurrences (Table 6).

Table 6: Percentage of Transitive Infinitival Complements/All Infinitival Complements
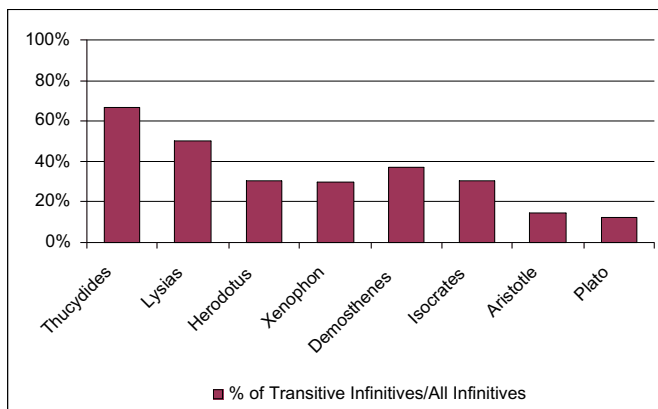
On average throughout the data set, only a fairly small proportion (approximately 20%) of periphrastic causatives have transitive infinitive complements.

Similarly, Greek allows the formation of causatives to verbs expressing action (*to flee*, etc.) as well as change of state (*to grow warm*, etc.) as demonstrated by the passages below from Xenophon's *Hellenica* and *Anabasis* respectively:

(9)   ὅλον ἐποίησε φεύγειν τὸ τῶν ἐναντίων (Xen. *Hell.* 7.5.24)
      'he caused the entirety of enemies to flee'

(10) θέρμῃ γὰρ λύειν δεῖ, ἢ ποιήσει ἀλεάζειν (Arist. *Pr.* 863b)
      'it is necessary to destroy (the cold) via heat, which makes (the cold) grow warm'

Table 7: Percentage of Action Infinitives/All Infinitives
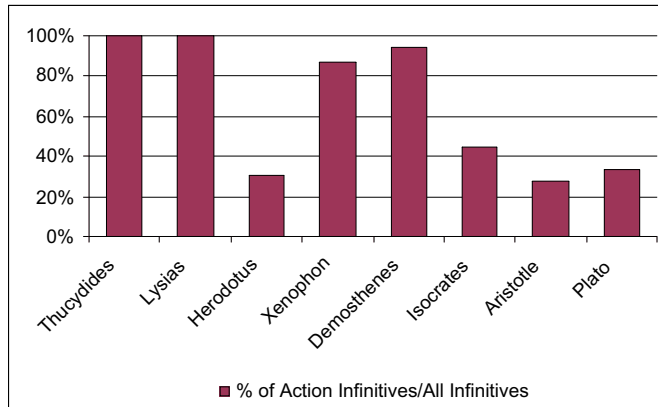
As might be expected, 'change of state' infinitives are much more common in Aristotle and Plato than in the other authors in the data set, where infinitives denoting 'action' are more frequent. Surprisingly, both Herodotus and Isocrates also have a very small number of action infinitives. In all of these authors, however, the large proportion of change of state infinitives is due largely to the frequent use of ποιέω with a copular verb such as εἶναι and a predicate adjective to express change of state in an inchoative sense (such constructions are mentioned frequently in the standard grammars: e.g. Kühner-Gerth 1966: 2.2.28).

(11) τὸ ἀρχαῖον ῥέεθρον διαβατὸν εἶναι ἐποίησε (Hdt. 1.191)
      'he made the former stream to be crossable'

These constructions resemble periphrastic factitives with copular verb elided (compare English *make X red* vs. *make X to be/become red*).

3.5. So far we have seen that, outside of Plato and Aristotle, the causer more frequently motivates the causee to do some action rather than change state. We will see now that the semantic relationship between the causer and causee is markedly different in Aristotle and Plato as well. Ancient Greek causative constructions with ποιέω can express all four possible combinations of +/–Animate causer and causee, though some combinations occur more frequently than others. The distribution is as follows (Table 8):

Table 8: Animacy Relationship of Causer and Causee by Author
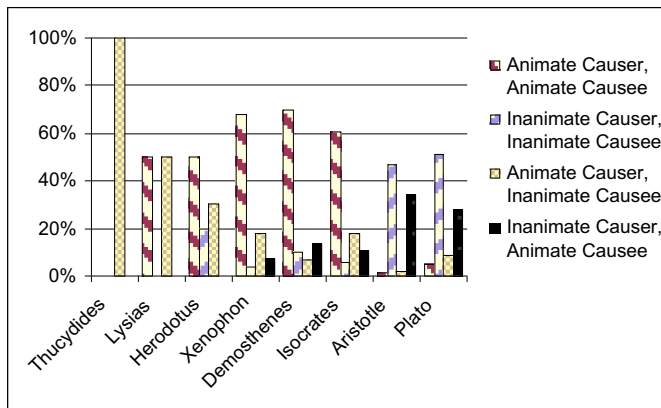
N.B. Totals do not add up to 100% in authors with objects that are not explicit.

3.5.1. The first possible combination is a –Animate causer and –Animate cause. This combination occurs most frequently in Aristotle and Plato, but quite rarely in other authors. For the most part, these represent very abstract instances of natural cause, as in example (9) above from Aristotle, in which heat causes the cold to warm up. The object of the construction is inanimate and therefore incapable of being 'unwilling'.

3.5.2. The second possible combination has a +Animate causer and a +Animate causee. This is the most frequent type of causative with ποιέω in all authors besides Plato and Aristotle. (Thucydides appears anomalous, but the size of the data set (only 3 causatives) is too small to draw firm conclusions.) Causatives with +Animate causer/+Animate causee parameters have semantics that come very close to constructions with κελεύω and ἀναγκάζω + infinitive, where a person causes others to commit some action which they are not necessarily willing or inclined to do:

(12) Ῥᾴδιον μὲν οὖν ἐστί μοι παῦσαι τῆς βλασφημίας αὐτοὺς μικρὰ πάνυ προεμένῳ, καὶ ποιῆσαι λέγειν ἐπαίνους ὑπὲρ ἡμῶν. (Dem. *Ep.* 20)
'Now it is easy for me to make them stop their abuse, at a very small expense, and make them to speak praises about us.'

In the passage above from Demosthenes, the element of verbal command or compulsion which would be implied by using κελεύω and ἀναγκάζω is for whatever reason not desired — perhaps because of the more subtle suggestion of bribery.

3.5.3. The third possible combination has a causer which is +Animate but a causee which is –Animate; this type is quite rare in all authors.

(13) καὶ ἀντὶ Σικανίας Σικελίαν τὴν νῆσον ἐποίησαν καλεῖσθαι (Thuc. 6.2)
       '(the Sicels) caused the island to be called Sicily instead of Sicania'

In many of the constructions of this third type, the causee is a legal or political institution that might properly be considered animate via personification, in which case this category could be subsumed almost entirely under the second category above, causatives with a +Animate Causer and a +Animate Causee:

(14) ἔπειτα ὅτι τὰς ναῦς προύδωκε καὶ τὴν πόλιν περὶ σωτηρίας βουλεύεσθαι
       πεποίηκεν (Lys. 26.23-4)
       'then, the fact that he gave over our ships and made the city consider its safety'

(15) καὶ ταύτην τὴν βουλὴν συνδικάζειν ἐποίησεν ἐν ᾗ Σάτυρος μὲν καὶ Χρέμων
       μέγιστον ἐδύναντο (Lys. 30.14)
       'and that he made that same council to serve as assessors at the trial in which Satyros
       and Chremon had the most power'

3.5.4. The fourth and final category of constructions have a causer which is –Animate but a causee which is +Animate. Such constructions are rare outside of Aristotle:

(16) τὸ μὲν γὰρ μικρὸν προσπταίειν πολλάκις ποεῖ τὸν ἀκροατήν ... τὰ δὲ μακρὰ
       ἀπολείπεσθαι ποιεῖ (Arist. *Rh.* 1409b)
       'the small clauses often cause the hearer to stumble while the large ones cause him to
       get left behind'

In constructions where the subject is inanimate, we can assume that verbs requiring an animate subject like κελεύω or ἀναγκάζω would be impossible, whereas ποιέω is perfectly acceptable.

4. To summarize, the 497 attested examples of causative constructions in Greek prose from Thucydides to Aristotle show that, in contrast to the *communis opinio*, Greek had grammaticalized the causative construction with ποιέω well before the *koine* period. While the prevalence of examples in Plato and Aristotle might suggest that philosophical language, with its need to discuss causal events, was reponsible for the creation and establishment of periphrastic causatives, this view is clearly wrong on several grounds. Syntactically, the

construction is as early as Homer (e.g. *Odyssey* 23.258 above), where the shift in meaning of ποιέω from concrete 'make' to 'cause' had already taken place well before the two philosophers. While at best Plato and Aristotle may have popularized the construction, they cannot have created it. Even the view that the philosophers popularized the construction is unlikely to be true: first, if philosophical language had played an important part in the diffusion of ποιέω causatives, we would expect the constructions to continue to be found in intellectual or cultured rather than colloquial registers. To the contrary, in New Testament Greek, ποιέω causatives occur most frequently in simpler prose style (i.e. *Revelation*); as alluded to above, the earlier verse literature seems to confirm the colloquial nature of the causatives. For example, in Aristophanes we find the following threat:

(17) τὸν σησαμοῦνθ' ὃ κατέφαγες, τοῦτον χεσεῖν ποιήσω (Ar. *Thesm.* 570)
      'I will make you crap forth the sesame cake that you have eaten'

Furthermore, as we have seen, the causative use found in Plato and Aristotle has a number of peculiarities of its own (frequent inanimate subjects, lack of explicit objects, large numbers of intransitive infinitives denoting 'change of state', etc.) which are not matched, at least statistically, in other authors. The high frequency of periphrastic causatives in Isocrates suggests that the construction was generally widespread outside of philosophy. The hypothesis of a colloquial origin for ποιέω plus accusative and infinitive requires further demonstration on the basis of more extensive evidence. We may at least conclude here that in the course of the late 5th century and early 4th century, the causative constructions developed as part of the standard language until philosophers adopted them and extended their use in abstract language to which they did not originally belong.

References

Chantraine, P. (1991). *Morphologie historique du grec*. Paris.
Cooper, G. L. (1998). *Attic Greek Prose Syntax*. Ann Arbor.
Dixon, R. M. W., and Aikhenvald, A. (2000). *Changing Valency: Case Studies in Transitivity*. Cambridge.
Humbert, J. (1960). *Syntaxe grecque*. Paris.
Kühne, W. (1882). *Das Causativum in der griechischen Sprache*. Leipzig.
Kühner, R., and Gerth., B. (1966). *Ausführliche Grammatik der griechischen Sprache*. Hannover.
Meier-Brügger, M. (1992). *Griechische Sprachwissenschaft.* Berlin.

Meillet, A. (1965). *Aperçu d'une histoire de la langue grecque.* Paris.

Moreno, J. C. (1993). '"Make" and the semantic origins of causativity: a typological study', in B. Comrie and M. Polinsky (eds.), *Causatives and Transitivity*. Cambridge, 155-64.

Rix, H. (1992). *Historische Grammatik des Griechischen.* Darmstadt.

Schwyzer, E., and Debrunner, A. (1939-71). *Griechische Grammatik.* Munich.

Song, J. J. (1996). *Causatives and Causation: A Universal-Typological Perspective*. London.

Wackernagel, J. (1926-8). *Vorlesungen über Syntax*. Basel.

Simplification of case morphology in Koine Greek

Evidence from the private letters from Mons Claudianus

Ela Harrison[*]

1. Introduction

In this paper, I study the language of the private letters (nos. 137-75, 177-9, 220-303 as published by Bingen *et al.* 1992, 1997) written between *c*. 110-160 AD at Mons Claudianus in the Eastern desert of Egypt. These non-literary texts show live evidence for the restructuring of the case system of Greek known to have taken place between the late Koine period and the early Mediaeval period. In particular, they show advanced ways in which fixity of word order is taking over from morphological marking as a means of expressing linguistic relationships. In Section 2, I describe the physical background of the texts and their authors. In Section 3, I state, with brief exemplification, that the orthography and morphology of the letters is in line with that of contemporary non-literary texts. This concurrence with comparable texts establishes that generalizations about the state of the language made on the basis of these letters may have wider validity. In Section 4, I show that in this special pragmatic context of private letters, with formulaic expressions and world knowledge, fixity of word order takes over from case marking to express argument structure relations to a marked extent, and especially in the area of personal names.

2. Background

Mons Claudianus was peopled by a Roman army garrison (principally the Ala Apriana) and also quarry workers, merchants, and craftsmen. It lies in the Eastern desert, five days' camel journey from the Nile Valley — and is thus a long distance from the nearest inexpensive source of papyrus.[1] This explains why potsherds are used for lengthy private letters and even literary texts at Mons

---

[1] Papyri found at the site are both far fewer and far less well preserved than the ostraca.

Claudianus, whereas elsewhere they are generally only used for receipts, accounts, etc. Over 6,000 pieces of writing were found at Mons Claudianus, so that the letters, inventories, school exercises, etc. published to date (numbering around 400) represent a small sample of the total.[2]

Our letters range from a couple of lines to 20-odd lines long. Some are in perfect condition, others too fragmentary to establish continuous sense.[3] The subject matter of the Vol. I letters[4] is largely personal but also includes some transactions. Transactions form the bulk of the subject matter of the letters from Vol. II:[5] several letters evidently accompanied consignments of vegetables and others contain requests for supplies of various sorts. Most of the texts recovered at the site are in Greek; there are also Latin texts and very few texts in Demotic.[6]

3. Phonology and morphology

The phonology and morphology represented in the letters accords with that of other non-literary texts of this period as described in Gignac's accounts (1977, 1981). The letters show abundant evidence for the various sound changes undergone in this period. I show a few representative examples and remark on how these changes affected the morphological case system.

3.1. Phonology

Vowels are no longer distinguished for length. The loss of distinctive vowel length has resounding implications for the morphology, especially for the 2nd declension 'o-stems'. This loss is represented in the spelling by the exchange of 'long' and 'short' graphemes both in the word stem — e.g. Λωνγεῖνος (164) for Λογγῖνος —, and in case suffixes — e.g. Ζοσίμο (151) for Ζοσίμωι (with monophthongization of the original long diphthong in ωι), ἀδελφό (220, 222, 247, 262, 275) for ἀδελφῶι (see previous comment), and conversely τώ (235, 270, 279) for τό.

Iotacism. Several high and mid front vowels have merged into /i/ (cf. Gignac 1977: 189-91, 235-42, 272-3). Examples are χοιρειδείου (138) for χοιριδίου;

---

[2] I am grateful to Dr Walter Cockle for much helpful information on the find context and prosopography.

[3] Some of these latter are published as part of an archive in an identified hand, allowing reliable reconstruction.

[4] The letters of Vol. I come mainly from the south quarter of the excavation site.

[5] Most letters of Vol. II originate from locations secondary to the main camp.

[6] The question of whether the authors of the letters were native speakers of Greek is an important one for which I do not yet have a straightforward answer. My impression from lengthy and close study of the texts is that many, but not all, were native speakers.

Σεραπιάδει (138) for Σεραπιάδι, ἱππῖ (177) for ἱππεῖ, κιβαριάτι (244) for κιβαριάτηι. These last three are datives from different inflectional stem types but the endings are now pronounced identically, so the formal distinction between the 3rd declension consonant stems (Σεραπιάς), /eu/ stems (ἱππεύς) and the 1st declension (κιβαριάτης) is undermined.

The sounds originally represented by <υ>, <υι>, and <οι> are confused in these letters and elsewhere in the Greek of Egypt (cf. Gignac 1977: 197-202, 293-4); e.g. Χύακ (260), for the Egyptian month name Χοίακ, πυήσις, πυῶ, etc. for ποιήσεις, ποιῶ (passim e.g. 153, 243 x 2), τῦς for the article τοῖς (263), συ (227, 245, 265, 278) for σοι.

The diphthongs of Classical Greek have monophthongized (Gignac 1977: 191-7); e.g. Εἰρενέο (273) for Ἰρεναίωι, δέσμε (283) for δέσμαι, ἄδελφαι (237) for ἄδελφε, μαι (270) for με, σαι (261) for σε.

The important point to be noted is that these changes affect all areas of the vocabulary — lexical words including loanwords, function words, personal names, pronouns, and numerals.

3.2. Morphology of the nominal system

The morphology is standard for the late Koine period as documented by Gignac (1981). Two sorts of movement towards simplification of the morphological system are particularly evident. Third declension nouns, especially morphologically irregular ones such as s-stems and r/n-stems, tend to be replaced by a 1st or 2nd declension synonym (e.g. ὀψάριον[7] replaces ἰχθύς) or by an alternate form with a 1st/2nd declension diminutive suffix (e.g. ἰχθύδιον (242) replaces ἰχθύς). As demonstrated, both processes may be used for the same word. On the other hand, a large number of unusual terminations are in use solely for personal names, mainly those of Egyptian origin (e.g. in -ους, -ων, -ως, -ας).

The Koine's wealth of productive suffixes and derivational compounds[8] is amply demonstrated. Especially widespread is the (originally diminutive) -ιος/-ιον suffix. A large stock of words (and of personal names and Latin loanwords) with this suffix underwent apocopation of the /o/ to form a whole new sub-declension of the second declension in -ις/-ιν (cf. Gignac 1981: 25-9,

---

[7] ψαρίδια (225), with apocope and the diminutive suffix, is exactly the modern form.

[8] Some derived forms are unattested elsewhere, such as the compound ὁλοχάλκινεν (279) and the diminutive ῥακάδιον (174); others go back to the 4th century, such as κίθωνιν (with metathesis), οἰνάριον (cf. Gignac 1977: 93-4).

115).[9] Whilst the many loanwords from Latin are fully incorporated into Greek morphology, the few Egyptian loanwords (just one or two words, e.g. *taskou* — a kind of jar) are invariant.

4. The reorganization of the case system

The disappearance of the dative case, its partial syncretism with the genitive and take-over by prepositional phrases has been extensively studied (e.g. Humbert 1930, Kurzova 1986, Luraghi 1986, Horrocks 1997). However, the exact means by which this reorganization was effected still requires clarification. I shall show below, with reference to three specific areas, that at this stage the genitive case is discernibly taking over the dative case functions and losing its partitive valency. Moreover, in each of the three contexts, fixed word order is demonstrably 'filling in' for the morphological expressiveness typical of the standard language.

4.1. Proper names

194 names are attested in these letters.[10] Between them, these names represent a far larger set of inflectional stems than is employed for the regular nouns. In addition to the Roman names, which, like the Latin loanwords in general, are easily assimilated to Greek inflectional stems, there are many Semitic names with their own idiosyncratic stems (e.g. in *-ous -outis*, *-as -atis*). However, morphological anomaly alone is not the deciding issue where unexpected case marking is attested in the area of proper names. For example, the town name *Raima* is never inflected but could perfectly easily be declined as an a-stem. Similarly, the names *Eremisis* and *Pouonsis* are invariable, but names like *Orsenouphis* and *Psenpaapis* do get inflected. In what follows, I focus on the personal names as opposed to toponyms or month names (there are attested the month names *Hathyr*, *Choiak*, *Phaophi*, all invariable, and the month name *Pachon*, which is abbreviated). I discuss the three main contexts in which personal names are attested and show that the fixity of the word order in each of these contexts 'does duty' for the case marking where names are concerned.

---

[9] In the letters, the full -ιος/-ιον and the new -ις/-ιν are both in evidence side by side in apparently free variation. If anything, the tendency for Latin names to take -ις/-ιν seems particularly strong.

[10] I include only those names whose endings are *preserved* rather than reconstructed in this count.

4.1.1. 'Vital statistics'

Of the194 names attested, 92 appear only once, in the expected case. 23 appear more than once but are only attested in one case — the expected one.[11] 12 only appear in one case, not the appropriate one. These are *Artemidorus*, *Besarion*, *Chennamis*, *Eirinaios*, *Furius*, *Kol*, *Kopres*, *Laberas*, *Maronas*, *Taesis*, *Tithoes*, plus [ ]εμων (apparently nominative case for expected dative recipient). Two appear in a fragmentary context such that it is impossible to tell whether they are in the 'right' case or not. The name *Eutuchos* appears twice but both times it is abbreviated so that no case ending is attested. Of the 57 names that appear in more than one case, 31 are always assigned the expected case whilst 26 are assigned the 'wrong' case at least once. These are *Ammonianus*, *Ammonius*, *Claudi(u)s*, *Dionysios*, *Dioskoros*, *Drakon*, *Eponychos*, *Eremisis*, *Herakleides*, *Herminus*, *Iano(u)aris*, *Isis*, *Leon*, *Longinus*, *Marinus*, *Maximus*, *Nemesion*, *Orsenouphis*, *Paniskos*, *Petoseris*, *Pouonsis*, *Psenosiris*, *Sarapammon*, *Serenus*, *Titiois*, *Tyche*. The personal names that get assigned an inappropriate case represent Greek, Roman, and Egyptian names and in fact constitute a reasonably representative sample of the mix of names attested at Mons Claudianus, which speaks against an explanation based on the unfamiliar stems of foreign names.

4.1.2. Opening formula

Personal and official letters employed various formulae which were almost invariably used. 110 of the 124 letters preserve the standard opening formula 'NN(s) (adjectives) (*expected case: Nominative*) to NN(s) (adjectives) (*expected case: Dative*), greetings'. 79 have the expected cases in the opening formula. In 14 letters, only one of the names is preserved. 17 letters contain an unexpected case inflection in the greeting formula: the archive of letters from Dioskoros (224-34) account for 8 of these. Fifteen involve a name of the recipient of the letter in the nominative case instead of the expected dative case: these are the ones taken as significant here.[12]

(1)   Διόσκορος Δράκων καὶ Ἐρεμησις καὶ Ἀμμωνιανὸς κουράτωρ καὶ Πετοσερις καὶ Πανίσκος ἀμφοτέροις τοῖς φιλτάτοις…

---

[11] This includes the two instances of Μαύρο for Μαύρωι, counting the form as a spelling mistake rather than incorrect case choice.

[12] The other two instances (letters 255 and 256), both from Livianos, both involve the form φιλτατε for expected dative, construing with a name in the dative case as expected. See Bingen's notes *ad loc.*: this may be reduction of /o/ to /e/ in unstressed syllable rather than incorrect case assignment. The same comment applies to the form Επονυχε for the sender of letter 238. Letter 238 is in the same hand as the archive of letters from Dioskoros.

> 'Dioskoros to Drakon and Eremisis and Ammonianus the *curator* and Petoseris and Paniskos, 'both'[13] good friends…' (225)

There are several letters from Dioskoros to this group of people, as well as letters from other authors to the same group. We know that Dioskoros is the sole sender and that the rest of the names designate recipients because of the external evidence of the archive, because Drakon is explicitly addressed in person in the course of some of the letters, and because of the word order. Where there is more than one sender, the two names would always be linked with καί.

A striking example of the fixity of word order taking over from case marking to express the sender-addressee relationship in the context is the following. Here, the authors' names are in the dative and the recipients' names in the nominative but the adjectives in apposition to the second pair of names are 'correctly' in the dative:

(2)   Ἡρακλείδηι καὶ Διονυσίωι Πανίσκος καὶ Ἑρμῖνος τοῖς ἀμφοτέροις…
   'Heraklides and Dionysios to both Paniskos and Herminos…' (280)[14]

The fact that the *senders'* names are in the dative in this example shows that case marking is not the mechanism by which the sender-addressee relationship is expressed but that the word order does this job. There is an interesting asymmetry here between the personal names of the recipients in the nominative case and the definite article + adjective τοῖς ἀμφοτέροις in the dative as expected. Given that *Paniskos* and *Herminos* are simply o-stems, morphological oddity is clearly not the sole explanation.

4.1.3. Personal names as complements of prepositional phrases

There are 20 instances of post-prepositional names being either in the nominative or else assigned an inappropriate oblique case.[15] The total count of post-prepositional names in the letters is 70, so this is a widespread phenomenon.

Three of these instances are from *proskynema* formulae — 'I do your *proskynema* (every day) by (= παρά) (deity)', so the name is that of a deity

---

[13] The adjective ἀφμόφτερος 'both' is part of the letter formula and is used in these letters and elsewhere where there are more than two addressees as well as when there are exactly two.

[14] This ostracon was found with — and comes from the same amphora as — no. 279, which opens Ἡρακλείδης Πανίσκος τῶι φιλτάτωι πολλὰ χαίρειν. This helps with the belief that Herakleides was the sender of 280 and not a recipient.

[15] The place name *Raima* is also found in its invariant form five times following prepositions and the month names *Hathyr* and *Choiak* once each.

rather than a person. The name is genitive for expected dative twice and nominative for expected dative in the third instance.

(3)   Τὸ προσκύνημά σου ποιο παρὰ τῆς κυρίας Ἴσιδος
       'I make your obeisance every day by the lady Isis' (268)

The other 17 instances have to do with the quasi-formulaic locution 'receive X from NN (the *familiarius/*camel-driver/wagoner)'. 13 of these have the name in the nominative; two, both from the same hand, have dative for expected genitive; one has the name abbreviated and one has the name in the accusative for expected genitive.

(4)   κόμισαι παρὰ Λογγινᾶτι…
       'Receive from Longinas…' (249) (dative for expected genitive)

(5)   κόμισαι παρὰ Κωλ τὸν ἁμαξέα...
       'Receive from Kol (nom.) the wagoner (acc.)...' (177)[16]

Whilst Classical Greek uses polysemous prepositions whose meaning is made explicit by the case with which they construe (cf. Luraghi 1986), Modern Greek has unambiguous prepositions followed invariably by the accusative.

In the *proskynema* formula the preposition used is παρά + (expected) dative, to indicate the location of the worship, and in the 'receive from NN' formula it is παρά + (expected) genitive, indicating the source. Aside from the tendency to have personal names in the nominative in these formulaic contexts, the genitive and the dative are being substituted for one another after a preposition, so that case marking alone no longer disambiguates the polysemous preposition. Rather, the preposition must indicate by itself whether source or location is indicated. As with the opening formula, the word order crucially determines the dependence of the personal name on the preposition, not the case marking. Once again, there is an interesting asymmetry between personal names and regular nouns in this.

---

[16] Note that the noun in apposition is in the *accusative*, rather than expected genitive; cf. διὰ Λαβῆρα (accusative for expected genitive in a personal name) (292) — see below. We do not know how the name Κωλ would have been declined.

4.1.4. Sundry greetings

Of the 25 letters containing requests to pass on greetings to other named individuals, 6 (all from Vol. 2) have the names of the individuals to be greeted in the nominative rather than the expected accusative:

  (6)   ἀσπάζομαι Ὀρσενοῦφις καὶ Μαρῖνος καὶ Λέων
         'I greet Orsenouphis and Marinos and Leon' (258)

  (7)   ἀσπάζομαι τὸν φίλον σου Ἑρμῖνος καὶ Κλαύδις Ἐπώνυχος
         'I greet your friend Herminos and Claudius Eponychos' (279)[17]

This is the third main context in which personal names occur in these letters. Once again, the form of words and the word order are fixed so that it is clear that the named individuals are the intended recipients of the greetings despite the fact that they are in the nominative case. Once again (ex. 7) there is an asymmetry between personal names in the nominative and regular nouns in the expected accusative.

A parallel that will suggest itself to many for this phenomenon of having personal names in the nominative case instead of an expected oblique case is the routine inflectional invariance of Semitic names in the Greek of the Septuagint and of the New Testament. (Similarly, Hebrew names and Latin names and loanwords in Old Irish, Finnish, and other languages are invariable.[18]) However, it is worth emphasizing that *by no means all the names affected are Semitic*.[19] Nor is this an idiosyncrasy of the Mons Claudianus writers. An examination of the ostraca published in Tait (1930)[20] and of the papyri in Davis (1932)[21] reveals that a striking number of these also exhibit personal names in the nominative for expected dative in the context of greeting formulae and prepositional phrases. This tendency seems to have been in train up to 300 years earlier than the letters currently under consideration.

---

[17] Here, the qualifying noun in apposition to the name gets inflection, as in 270 but not in 225 cited above. See 4.1.5. below for further discussion of how nouns in apposition to personal names are treated.

[18] I am grateful to Rolf Noyer and Paul Kiparsky for interesting discussions about this issue.

[19] Gignac (1981: 103) finds Egyptian and Semitic PNs frequently used indeclinably, especially in the Byzantine period (our data is of course much earlier), but also notes that most have alternative formations that can fit into Greek paradigms.

[20] These originate from various settlements in the Nile Valley and date from the 2nd century BC to the 2nd century AD.

[21] These from various locations in Egypt and mainly the 1st century BC.

4.1.5. Other lexical items in these contexts

I have remarked upon the asymmetry between personal names in the nominative and regular nouns in the 'expected' case in each of the three main contexts in which personal names are attested. I now take a closer look at the treatment of these regular nouns (also adjectives and a few participles) for the purposes of comparison. (In this connection, it should be repeated that the number of inflectional stems represented in the names is far greater than amongst common nouns.)

Of the 40 lexical-item types found in these contexts,[22] 17 appear just once, in the expected case. Two appear more than once but only in one case, the expected case. One appears just once, not in the expected case. Of the 20 words that recur and are used in more than one case, 15 are inflected as expected; 5 are sometimes not. There are some qualitative differences that should also be mentioned. Whilst the name Μαῦρος is twice written Μαύρο for dative, this exchange of <o> and <ωι> occurs 4 times in the 46 occurrences of τῶι ἀδελφῶι in apposition with the addressee and 4 times in the 12 occurrences of τῶι φιλτάτωι in the same context.[23] ἀμφοτέροις, which occurs 12 times, is also once spelled ἀμφοτέρο: this in apposition to names in the nominative for expected dative but prefaced by τοῖς τρισί in the expected dative. Thus, this 'spelling mistake' is far more frequent in the nouns and adjectives than in the names.

Of the 6 words that appear in the 'wrong' case, one appears in opening formulae, two in the 'sundry greetings' section and three in the 'receive x by way of y' formula. The explanations for these 6 are various. The title κουράτωρ (nominative for dative) appears in the opening formula in apposition to Ἀμμωνιανός (nominative for expected dative of addressee) in letters 224-38: Dioskoros' letters, which all have the names of the addressees in the nominative but ἀμφοτέροις τοῖς φιλτάτοις 'correctly' in the dative (cf. ex. 2). In these letters, it appears that the title κουράτωρ is treated as *part of Ammonianus' name*.

In the 'sundry greetings' sections of the letters, πάντες and φιλοῦντες are apparently nominative for accusative 5 times out of a total of 7 occurrences apiece. However, the article that goes with them each time is the expected accusative τούς. This demonstrates the morphological restructuring of the 3rd declension paradigm as continued into the modern language and is also instantiated in the other 3rd declension plurals seen in these letters, viz.

---

[22] The set of adjectives and appositive nouns used in these letters is small, indicative of the formulaic nature of the context of many occurrences of names.

[23] το αδελφο, το φιλτατο twice each; το αδελφωι, το φιλτατω, τωι αδελφο, τωι φιλτατο once each; το φιλτατε twice: cf. note 12.

στατῆρες and τέσσερες. This is thus not 'misinflection', but a morphological change in progress.

The remaining three words are all agentive nouns, appearing in apposition to personal names. Interestingly, these are not in the nominative like the names but are assigned an inappropriate case. τὸν ἁμαξέα (ex. 5) is accusative for genitive. καμηλίτης, which occurs 7 times, is 'correctly' inflected 4 times, abbreviated once and put in the dative instead of the genitive twice. φαμ(ε)ιλιάρι(κ)ος occurs 4 times, twice in the appropriate case, once in the dative instead of the genitive, and once in the genitive plural instead of the genitive singular. All of the instances where the incorrect case is assigned are in apposition with a personal name in the nominative case.[24] Thus, once again we are seeing confusion of the genitive and dative cases in the environment of a preposition.

4.2. Clitic pronouns[25]

The pronouns με/μου/μοι, σε/σου/σοι appear 292 times in the corpus. They are of interest in the present context because, given their status as enclitics, they display some fixity of word order by definition. Of the 292 instances, 16 are not in the expected case. The instances of inappropriate case assignment attest to the erosion of the dative case. The four 'inappropriate' accusatives are in place of an expected dative, as are the seven 'unexpected' genitives. There are also five 'unexpected' datives: four of them as direct objects of verbs and the fifth in the accusative + infinitive of a greeting formula. Whilst μοι is spelled as μυ 9 times, σοι as συ 6 times, and με, σε as μαι, σαι once each, there is no reason to believe that these pronouns were confused in pronunciation.

(8)   ...ἐγὼ πέμπω σε. πέμσον με τὸ πελέκειον
       'I (sc. will) send you (the pig). Send me the axe.' (acc. in place of dat. for indirect object twice) (138)

(9)   ...μὶ μέμψε μαι ὅτι οὐκ ἔπεμψά σου λάχανο. ἠὰν γένεται, πέμψω σου.
       'Don't blame me that I didn't send you vegetables. If there are any, I'll send you.' (gen. in place of dat. for indirect object twice) (270)

---

[24] 12 corresponding forms with no personal name are all inflected as expected.

[25] Gignac (1981: 161) notes that the first and second person personal pronouns especially show morphological developments that are further advanced than those found in the nouns and adjectives and seems to imply (e.g. p. 163) that instances of 'misinflection' are not uncommon.

As is obvious from the fact that each of these examples contains two of the 'mistakes' and thus together they represent a quarter of the total count, only a few of the writers do this. As far as we can tell from the archives of individual writers, there is a certain amount of correlation between the tendencies not to inflect personal names and to assign an inappropriate case to clitic pronouns. Here again, the fixity of word order in formulaic contexts 'does duty' for the case marking in making argument structure relations clear. Thus

> (10)   πρὸ μὲν πάντων εὔχομέ *σοι* ὑγιαίνει(ν)
>   'First of all, I hope that you are well' (224)

has the object 'you', which is simultaneously the subject of the infinitive, in the dative instead of the accusative. This expression is so common in the letters that the case marking seems to be pro forma. Another example, with the plural pronoun, even has the verb ending malformed.

> (11)   πρὸ μὲν πάντων εὔχομα(ι) ὑμῶν ὑγιαίνων
>   'First of all, I hope that you are well' (?) (238)

This is from a letter where the names of the addressees are also in the nominative for expected dative.

4.3. Partitive genitives

Besides evidence of confusion between the genitive and dative cases, the decline in the use of the genitive in another of its core functions — the 'partitive' use — is also discernible. This is interesting because it suggests that as well as the changes in the case system tending towards the erosion of the dative case and its replacement by the accusative or genitive as indirect object, there is a shift in the core functions of the genitive itself.

Of 23 instances of partitive genitive constructions in these letters eight — a third — have the commodity in the nominative or accusative rather than expected genitive. It should be noted that in Modern Greek the partitive genitive is only found in formal contexts and the preferred locution is precisely this 'recipe style' apposition (cf. Mackridge *et al.* 1995: 268-9, 345-6).

> (12)   ...κομίσατε σεύτλια δέσμην γ καὶ ἄλλην δέσμην σέρις
>   'receive 3 bunch (sic) beets and another bunch chicory' (nom. for gen. twice) (228)

It should be noted that four of the instances are due to the Dioskoros writer (226, 227, 228, 238). However, this usage is not restricted to the Dioskoros writer nor to consignments of goods:

(13)   ...τὸ πρόσλοιπον τὴν τειμήν
          '…(send me) the rest the price' (acc. for gen.) (139)

(14)   ...καταγραφὶν γράφις τοὺς ἐργάτας
          '…write a list the workers' (acc. for gen.) (141)

Thus, the tendency towards the appositive locution that later wins out in the language can already be seen beside the partitive use of the genitive in the 2nd century AD. Once again, juxtaposition rather than inflection makes the relations clear.

5. Conclusions

The language of these letters has been shown to be representative of contemporary non-literary texts. Whilst examination of the texts reveals plentiful evidence for changes in the use of the case system — notably the erosion of the dative case and its eclipse by the genitive —, the most striking phenomenon is the way that the fixity of word order 'does duty for case marking' in the context of these letters with their repetitive subject matter and formulaic usage. As we have seen, this is especially so with personal names. Despite the fact that names comprise a larger set of inflectional classes than does the regular lexicon and include some morphologically intractable forms, it is clear from these letters that anomaly of morphological class is not the reason for inappropriate case assignment. The letters provide a context in which formulaic material and world knowledge allow for a greater fixity of word order than is general. On reflection it is not entirely surprising that this is the locus of change and it may be possible to speculate that the alteration of the case system did indeed start in pragmatically defined contexts. However, it is also striking that personal names in these contexts behave *differently from other lexical items*, a phenomenon which is not normally registered in the grammars, without clear morphological motivation. It would be possible to speculate that the alteration of a case system proceeds gradually through different parts of speech. That pronouns have different case usage and case forms from nouns is well known. It may well be that personal names too are treated differently. If so, this phenomenon would have to be added to the growing amount of evidence that the

diachronic development of personal names does not always coincide with that of other nouns (Mańczak 1995, Morpurgo Davies 2000, etc.).

References

Bingen, Jean, Bülow-Jacobsen, A., Cockle, W. E. H., Cuvigny, H., Rubinstein, L., Rengen, W. van (eds.) (1992). *Mons Claudianus Ostraca Graeca et Latina: Vol. I*. Cairo.

Bingen, Jean, Bülow-Jacobsen, A., Cockle, W. E. H., Cuvigny, H., Kayser, F., Rengen, W. van (eds.) (1997). *Mons Claudianus Ostraca Graeca et Latina: Vol. II*. Cairo.

Davis, W. Hersey (1933). *Greek Papyri of the First Century*. New York.

Gignac, Francis Thomas (1977). *A Grammar of the Greek Papyri of the Roman and Byzantine Periods. Vol. I: Phonology*. Milano.

— (1981). *A Grammar of the Greek Papyri of the Roman and Byzantine Periods. Vol. II: Morphology*. Milano.

Horrocks, Geoffrey (1997). *Greek: A History of the Language and its Speakers*. London.

Humbert, Jean (1930). *La disparition du datif en grec*. Paris.

Kurzova, Helena (1986). 'Morphological semantics and syntax in the non-formalised sentence structure of Greek', in A. Rijksbaron, H. A. Mulder, G. C. Wakker (eds.), *In the Footsteps of Raphael Kühner*. Amsterdam, 147-59.

Luraghi, Silvia (1986). 'The opposition total : partitive and the use of cases with prepositions in Ancient Greek', in A. Rijksbaron, H. A. Mulder, G. C. Wakker (eds.), *In the Footsteps of Raphael Kühner*. Amsterdam, 177-92.

Mackridge, Peter, Holton, David, and Philippaki-Warburton, Irene (1995). *Greek: A Comprehensive Grammar of the Modern Language*. London.

Moulton, James Hope (1963). *Grammar of New Testament Greek. Vol. III: Syntax* (by N. Turner). Edinburgh.

Mańczak, Witold (1995). 'Morphologie des noms: règles de flexion, systèmes de flexion', in Ernst Eichler *et al*., *Namenforschung*. Berlin/New York, 427-31.

Morpurgo Davies, Anna (2000). 'Greek personal names and linguistic continuity', in Simon Hornblower and Elaine Matthews (eds.), *Greek Personal Names: Their Value as Evidence*. Oxford, 15-39.

Tait, J. G. (1930). *Greek Ostraca in the Bodleian Library and Other Collections. Vol. 1*. London.

'What name? What parentage?'

The classification of Greek names and the Elean corpus

Ina J. Hartmann[*]

Introduction

Names such as 'Horse-Lover', 'Who-has-True-Glory', 'Gift-of-God', 'Friend-of-Strangers', 'Ruling-through-War', and 'Having-Beautiful-Horses' evoke memories of American Indian names like 'Many-Horses' and 'Sitting-Bull'. In fact, the first five names listed above are typical examples of one of the two large categories of names found in Greek onomastics: the *Vollnamen* or 'compound names'. They sound much more familiar in their untranslated versions: Φίλιππος, Ἐτεοκλῆς, Θεόδωρος, Φιλόξενος, Πολέμαρχος, and Κάλλιππος.

1. The Greek Naming System

1.1. Compound Names (Vollnamen)

Compound names of the type Ἐτεοκλῆς are inherited from Indo-European. Similar names may be found in various IE languages, cf. e.g. Germanic *Burkhart, Hermann, Kuonrat (= Konrad), Friedrich*; Sanskrit: *Avanti-varman*, *Dāma-yantī* (see e.g. Solmsen 1922: 153ff.). They are formed from two (rarely more) complete lexical items. These items, called 'elements' in what follows,

---

tend to have an 'ehrenvolle Bedeutung' (von Kamptz 1982: 6; cf. also e.g. Solmsen 1922: 113, Masson 1995: 706). Further examples of typical compound names are Ἀγαθοκλῆς, Νικόμαχος, Ἵππαρχος, Ἀριστόδημος, Πατροκλῆς, Τιμηκράτης, Θεόδοτος, and Διόδωρος (cf. Solmsen 1922: 113-16).

There were very few restrictions on the formation of compound names: nearly any two of these 'ehrenvolle' elements may be combined with one another to create a new name. A good example of this combination of elements is the story of how Φειδιππίδης in Aristophanes' *Clouds* 60ff. is named by way of the combination of the different elements his father and mother respectively wished to include in his name (Solmsen 1922: 117; Masson 1995: 707).[1]

While a number of compound names are lexically meaningful, there are also names devoid of sense. This may arise on the one hand because of the very lack of constraints on name formation, and on the other hand because of the typical phenomenon of names developing into purely phonetic labels.[2]

1.2. Kurz- & Kosenamen

The second large category of Greek names are the *Kurz- & Kosenamen*, or simple and abbreviated names, such as Ζεῦξις or Πάτροκλος. When Fick introduced the term, he was convinced that almost all 'short' names were derived from *Vollnamen* and therefore used the terms *Kurznamen* and *Kosenamen* indiscriminately (Fick 1894: 15, 32-4). The derivational origin is easy to see in those *Kurznamen* which still show two stems, such as Κλέομμις, Μενέστας, Πάτροκλος. These are derived from Κλεομένης, Μενέστρατος, and Πατροκλῆς respectively.

The origin is less clear in Ζεῦξις, Ἀρίστυλλος and Ξανθώ (from Ζεύξιππος, Ἀριστοκλῆς, and Ξανθίππη). We can nevertheless be certain that derivation existed for both types because there are a number of cases where the *Kurzname* and full name are applied to one and the same individual. All these instances of 'Personenidentität' of mythological and historical persons have been carefully collected by Schmitt (1995: 706-7; cf. already Meister 1890: 173-4).

However, for the names built on a single stem, it soon became apparent that interpreting them in each case as abbreviations is an oversimplification. Bechtel, in his *Spitznamen*, took the important step of recognizing the names derived from sobriquets as a separate category (Bechtel 1898). These names tend to be

---

[1] Von Kamptz (1982: 9) points out that in contrast to what Solmsen (1922: 117) claims, the name need not be devoid of sense. The name Φείδιππος, on which it is based, is attested in Homer and matches the phrase ἵππων φειδόμενος.

[2] In the same way, in modern Italian a dark-haired, dark-eyed girl may be called *Bianca*, or an English girl be named *April* even if she was born in September. See also below note 5.

disparaging in nature, and often refer to flaws or to character traits, though reference to ethnic origin is common, too: Στράβων, Σίμων, Σύρος, etc. Names of this type are likely to have lost their etymological value, and to have become name-tags devoid of lexical meaning.[3]

In his *Personennamen*, Bechtel grouped these names, together with those derived e.g. from natural phenomena, such as Λύκος and Ἄνθος, under the unfortunate label 'die übrigen Namen' (Bechtel 1917: 477ff.). He then classified these 'other names' in eleven semantic categories, ranging from 'Menschentum und Lebensalter' to 'Der Name enthält eine Metonymie'. Bechtel acknowledged that it was not always easy to distinguish between the two groups, but Solmsen (1922: 128ff.) went further. He recognized that there were also 'von jeher einstämmige' names with an 'ehrenvoll' meaning, and made the important point that names such as Ἀγαθός do not need to be abbreviated, but could be names in their own right. Fraenkel (1935), too, used the phrase 'von jeher einstämmige Namen', and finally von Kamptz (1982), Masson (1995), and Schmitt (1995*a*) distinguished between *Kurz-* and *Kosenamen*, as well as between mono- and dithematic names.[4]

The result of this complex history is a very confused and confusing terminology which keeps being changed, adding to the general difficulties.

1.3. The different terminology of various scholars

As Table 1 shows, the differences are slight for the *Vollnamen* (Ἱπποκράτης) and the mono- and dithematic abbreviated names (Κλέομμις and Στράτιος). The first disagreements arise from the classification of Εὔοδος. Bechtel explains the difficulties concerning the classification of names formed from established dithematic words. He decides on calling these 'Vollnamen zweiter Schicht':

> An diesen Grundstock [der Composita, die eigens zum Zwecke der Namengebung vollzogen wurden] hat sich eine zweite Gruppe von Vollnamen angeschlossen, Composita, die nicht zum Zwecke der Namengebung geschaffen sondern von Haus aus Appellativa sind, die vermöge ihres Inhalts als Personennamen geeignet erschienen sind und, in die Function als Namen eingerückt, vermöge ihrer Doppelstämmigkeit das Aussehen von Vollnamen gewähren. (Bechtel 1917: VIII-IX)

---

[3] If Bechtel's interpretation of Πλάτων is right, this is exactly what happened to that name even before the philosopher was given it (Bechtel 1898: 6, also e.g. Notopoulos 1939: 135ff. and Davies 1971: 333), but the question of Plato's name is still hotly disputed.

[4] 'Monothematic' and 'dithematic' is the English terminology. Von Kamptz and Schmitt use 'einstämmig' and 'zweistämmig', while Masson contrasts 'simple' and 'composé'.

This interpretation seems to be accepted by most scholars, even though they may not say so explicitly. Von Kamptz, however, rejects these names as being too colloquial:

> Die Bedeutung [der Vollnamen] ist durchweg ehrenhaft und hält sich den festumrissenen Termini der Umgangssprache fern. Komponierte Namen, denen solche zugrunde liegen [...], sind daher nicht als V[oll]N[amen] anzuerkennen. (von Kamptz 1982: 6)

He classes Bechtel's 'Vollnamen zweiter Schicht' as *Kurznamen*, stressing that in Homer their number is not very high.

Confusion becomes much more acute in the last three columns of Table 1. Firstly, the terms *Kurzname*, *Kosename*, and *Kurzform* are used in different ways. Secondly, a number of names cannot be classified with absolute certainty (cf. Solmsen and Fraenkel 1922: 128, Bechtel 1917: X ff., and von Kamptz 1982: 15ff.). The differentiation is clearest in von Kamptz (1982: 10ff.), owing to his distinction of *Kurzformen* from *Kurznamen*. *Kurzformen* indicate abbreviated names, while *Kurznamen* are originally simple names (including compound appellatives, however), but the problems of an accepted classification remain.

Schmitt (1995*a*: 425) distinguishes between monothematic and dithematic names, and between *Kurznamen* and *Kosenamen*. He regards both the latter types as abbreviations from *Vollnamen*, but believes that *Kosenamen* possess hypocoristic suffixes, while *Kurznamen* do not. He also uses the term *Vollname* for names developed from sobriquets and for names from metonymics. However, his terminology is confusing because of its earlier connotations in other authors.

Part of these difficulties are due to the historical approach taken by all scholars. This requires a categorization based on the origin of the individual name; yet the names yield no such information. Solmsen comments on this with reference to names such as Ἀγάθων and Λύκος:

> Die herrschende Ansicht betrachtet diese als durch Verkürzung entstanden. Irgendein durchschlagender Grund dafür ist jedoch nicht vorhanden, da die einstämmigen Namen schon als solche einen befriedigenden Sinn ergeben und ihre Bildungsweise durchaus im Einklange mit den für Ableitungen von einfachen Wörtern geltenden Regeln steht. (Solmsen 1922: 128)

Scholar \ Example	Ἱπποκράτης	Εὔοδος	Κλεόμ(μ)ις	Στράτιος	Ἀγάθων	Στράβων	Λύκος
Fick (1894)	Vollname	Vollname	Zweistämmiger Kosename	Einstämmiger Kurz- & Kosename	Einstämmiger Kurz- & Kosename	Spitzname	Einstämmiger Kosename (oder Name aus Name?)
Bechtel (1917)	Vollname	Vollname zweiter Schicht	Zweistämmiger Kosename	Einstämmiger Kosename	Einstämmiger Kosename	Übriger Name	Übriger Name?
Solmsen (1922)	Vollname	Vollname	Zweistämmiger Kurz-(Kose)name aus Vollname	Einstämmiger Kurz-(Kose)name aus Vollname mit Gemination	Evtl. von jeher einstämmiger Kurzname ('ehrenvoll')	Evtl. von jeher einstämmiger Kurzname aus Spitz- oder Beiname ('tadelnd')	Evtl. von jeher einstämmiger Kurzname ('ehrenvoll'?)
Fraenkel (1935)	Vollname	Vollname?	Zweistämmiger Kosename (aus Kürzung hervorgegangen)	Einstämmiger Kosename ('ehrenvoller Sinn')	Evtl. von jeher einstämmiger Kurzname ('ehrenvoller Sinn')	Kurzname aus Spitzname ('tadelnd')	Evtl. von jeher einstämmiger Kurzname (Tier (ehrenvoll?))
von Kamptz (1982)	Vollname	Kurzname	Zweistämmige Kurzform	Einstämmige Kurzform	Kurzname oder Kurzform	Kurzname	Kurzname?
Dobias and Dubois (1990)	Nom composé	Nom composé	Diminutif avec deux éléments	Diminutif avec un élément	Diminutif avec un élément	Sobriquet	Surnom?
Masson (1995)	Composé	Composé	Diminutif du composé	Diminutif simple du composé	Diminutif simple du composé	Nom simple (sobriquet)	Surnom
Schmitt (1995a)	Vollname	Vollname	Zweistämmiger Kosename	Einstämmiger Kosename	Einstämmiger Kurzname	Einstämmiger Vollname	Vollname, urspr. Beiname

Table 1: Terminology of Greek onomastics according to various scholars

Four of the scholars in Table 1, Fick, Bechtel, Masson, and Schmitt, share the *communis opinio* in the case of Ἀγάθων, while the other four are uncertain whether or not to regard these names as abbreviations. For Λύκος, the consensus is much stronger: apart from Fick, nobody definitely classes the name as a *Kosename*.

1.4. An alternative categorization

As the difficulties noted above lie in the historical approach, it would be clearer and almost certainly more helpful to adopt a classification with no historical requirement. I therefore propose to classify all names purely on whether they are, or are not, compounds, without regard to the original form of the name. We would then have monothematic, i.e. non-compound names, on the one hand, and dithematic, i.e. compound names, on the other hand.

For a more detailed analysis, further distinctions could be introduced. Thus in the monothematic names, one could distinguish e.g. names without suffixes from names with suffixes: Ἀγαθός vs. Ἀγάθων.

For dithematic names, distinctions can be made between:

- full dithematic names: two complete, recognizable elements, e.g. Ἀριστοκλῆς.
- extended dithematic names: two complete elements plus suffixation, e.g. Ἀριστοκλείδης.
- abbreviated dithematic names: two elements, one of which is shortened, e.g. Ἀρίστοκλος.

In both categories, monothematic and dithematic names, one could distinguish between names which match lexical items (Ἁρμόδιος, Ἀβάσκαντος), and names that do not (Καλλίας). It remains to be decided whether the difference in accent between Πίστος and the adjective πιστός should be considered as equivalent to a lexical difference.[5]

---

[5] Note that a difference in the accentuation of names and nouns is not unique to Greek: e.g. in German there are some cases of unexpected accent shifts when lexical items are used as names. In all German words beginning with the prefix *ge-*, such as *Gemach*, *Gemahl*, *Gesang*, *gescheit*, *gewahr*, *gewitzt*, the accent falls automatically on the *second* syllable, as the Germanic prefix *ge-* never bears the accent. However, when the word *Gewalt* ('violence, force') is used as a surname, its actual pronunciation is ['Gevalt], with the accent on the *first* syllable. The reason for this shift lies, in my view, in the fact that names are labels. It would therefore be undesirable that this label, every time it is used, should provoke memories of the original meaning, especially if that meaning is, as in this case, entirely negative. To avoid the semantic connotation, the accent is therefore shifted to an unexpected place. The name in question is thereby alienated from the original lexical item.

If we use these criteria, the names quoted in Table 1 would be categorized as follows (adding Ἀριστοκλείδης, as Table 1 shows no example of an extended dithematic name).

Monothematic		Dithematic		
No suffix	Suffix	Full	Extended	Abbreviated
Λύκος	Ἀγάθων, Στράβων, Στράτιος	Ἱπποκράτης, Εὔοδος	Ἀριστοκλείδης	Κλέομμις

Table 2: Non-historical classification of Greek Names

This classification shows a higher degree of objectivity and clarity than the previous ones, but it is in some sense less informative as it deliberately avoids the historical approach. However, it is useful as a starting point from which we may reach more definite conclusions about the longstanding problems mentioned above. For this purpose, the basic distinction between monothematic and dithematic names is sufficient.

2. *The relationship of monothematic and dithematic names*

Three views of the relationship of monothematic and dithematic names are taken in the literature:

(1) All names were originally dithematic, and all monothematic names are derived from them.
(2) All names were originally monothematic; dithematic names are formed later as extensions.
(3) There are three categories:
  • original monothematic names
  • original dithematic names
  • monothematic names derived from dithematic names.

The first view is held, with the addition that some nicknames may be originally monothematic, by Fick (1894), the last by e.g. Solmsen (1922), Fraenkel (1935), von Kamptz (1982), Masson (1995), and Schmitt (1995*a*). The second view is favoured by Pulgram (1947) and (1960), though both articles concentrate on the form of the original Indo-European names.

In what follows, I examine the two extreme positions, with the aim of finding some concrete evidence for or against either of them or both. My findings lead to the conclusion that the third, moderate position is the most realistic.

2.1. The corpus

 The basis of the present work is a corpus strictly defined both geographically and chronologically.[6] The corpus in question comprises the personal names of Elis, the northwestern province of the Peloponnese within the borders of the Roman province of Elis, thus including Triphylia in the south. In the period between *c.* 800 BC and AD 400, 675 different names are attested in Elis, 588 of which are Greek. 86 are Latin, while the remaining names are either of unknown origin, or attested so fragmentarily that no reasonable reconstruction is possible. The names are shared by *c.* 1060 individuals.

For the present research, only the Greek names are of interest. For these, the *Lexicon of Greek Personal Names*, Vol. III.A (Fraser and Matthews, Oxford 1997), has been an invaluable source of information. Of the 588 names, 276 are monothematic, while 312 are dithematic (46.9% vs. 53.1%). Dithematic names are thus somewhat more common than monothematic ones (cf. also section 2.4. and Table 15).

2.2. Name formation and classification of elements

In the totality of the Elean name corpus, 289 different elements are used for name formation. Of these, 119 occur only in monothematic names, 108 only in dithematic names, and 62 in both. In order to achieve as objective an analysis of the forms as possible, the segmentation employed by Bechtel (1917) was chosen for the most part, since his aims differ widely from those of the present work. Where Bechtel gives several forms, e.g. ἀρχε-, ἀρχη-, ἀρχι-, ἀρχο-, one was taken at random, in this case ἀρχε-. Where he gives forms as being only the last element, e.g. -τροφος, for the sake of comparison, a form without the -ς was used: e.g. τροφο-. The hyphen thus does *not* give an indication of whether the element was used in the first or the second part of a dithematic name. As it was felt to be important not to loose information in this levelling process, however, the exact realization of the elements in question is noted as well as the standardized form.[7] The elements can be compared to one another on both

---

[6] In order to allow more flexible and transparent research, the material collected was entered in a database designed by the author. All information and statistics on the Elean personal names are drawn from this database, version of 04.02.2002.

[7] For a name Ἀρχίππων, the elements would therefore be ἀρχε-, with the realization ἀρχ-, and ἱππο- with the realization ἱππ-, plus the suffix -ων.

grammatical and semantic grounds. It is not always easy to decide whether a stem is nominal or verbal; here, too, Bechtel's decisions were accepted; thus e.g. in the case of ''Αρχε-, [...] zu ἄρχω' (Bechtel 1917: 78), the classification in the present corpus is 'ἀρχε-: verb'. 'Unclassified' refers to those elements which were either abbreviated to such an extent that they cannot be identified anymore (e.g. κ-), or to elements for which no origin and meaning can be established, e.g. ὀλιχ- ('Ολιχίδας) and ὑρκ- (''Υρκος).[8]

Below, the eight grammatical categories of the name elements and the number of elements in each are listed (Table 3):

Grammatical category	Number of elements	Examples
Adjective	48	ἀγαθο-, κλεινε-, φιλο-
Adverb/Negative	2	ἀ-, τηλε-
Numeral	1	πρωτο-
Preposition	8	ἐπι-, κατα-
Proper name	33	'Απολλο-, 'Ηρα-
Substantive	151	ἀναξ-, αἰχμο-, κλεο-
Verb	36	κρατε-, μενε-, σπενδε-
Unclassified	10	βαργ-, ὀλιχ-, ὑρκ-

Table 3: Grammatical categories of elements

The semantic categories are more complicated. Bechtel's categories cannot be used as he classifies only names, not elements, and only 'die übrigen Namen' at that (see above). This means that the dithematic names and their elements are not attributed to a semantic class at all. In order to class all elements, I have therefore devised new categories. Table 4 gives the categories and the criteria used for classification. It is important to note that the grammatical classification has no bearing whatsoever on the semantic one, though the analysis below will show that certain connections are likely.

---

[8] In the case of the latter, the obvious references such as Bechtel (1917), Masson (1990), and LSJ with its supplement have of course been checked.

Semantic category	Criteria	Examples
Abstract concept	Usually non-physical concepts not fitting in any of the other categories	ἀνια-, θαν-, φρων-
Action	Elements denoting any sort of activity	βασκε-, δω-, δρομο-, λιβε-, πραξι-
Body part	Any element denoting a part of the human body, including the mind	κεφαλο-, μερο-, νους, στομα-
Geography	Elements derived from place names or indicating origin	Ἀλφειο-, Ἀττικο-, Ἑλληνο-
Kinship	Any element having to do with kinship and family in the widest sense	γενο-, μητρο-, πατρο-
Nature	Any element connected with nature, including animals, plants, sounds of nature, etc.	ἀνθο-, ἱππο-, λυκο-, κελαδο-, κορυμβο-, πιτθο-, σταχυ-
Object	Any sort of object that does not fit in either Nature or War/Politics	δωρο-, πυλο-, στεφανο-, χρυσο-
Person/Age	Elements indicating age or type of a person	ἀνδρα-, γειτο-, νεο-
Quality	Positive or negative characteristics or features of both humans and objects	αἰσχο-, ἀριστο-, θρασυ-, ἰφι-, καλλι-, ὀξυ-, πιστο-
Religion	Elements derived from the names of deities, heroes, sanctuaries, or other words with religious connotations	Ἀπελλο-, Ἀπολλο-, Διονυσιο-, θεο-, Ἰσι-, μουσο-, Ὀλυμπ-, σιλην-
Time	Denoting periods or units of time	ἐτε-, ἡμερο-
War/Politics	Any element expressing leadership, honour, glory, etc., but also weapons	ἀρχε-, αἰχμο-, κλεο-, κρατε-, πολεμο-, πολι-
Unclassified	No lexical meaning comparable to the above	1. ἀντι-, προ-, ἀγχι- 2. βαργ-, ὑρκ-

Table 4: Semantic categories and their criteria

These categories are of course subjective and often fuzzy, as are the criteria and the final classification of elements. Table 5 gives the number of elements falling into each category. Grammatically speaking, the elements used for name formation are mostly nouns, i.e. adjectives or substantives: they make up 68.9% of the elements. Verbs account for 12.5%, names for 11.4%.

From a semantic point of view, five categories prevail, namely (in order of frequency): Quality (19.1%), War/Politics (16.3%), Nature (14.2%), Action

(13.1%), and Religion (11.1%). Together, they account for 73.8% of the elements.

Semantic category	Number of elements
Abstract concept	6
Action	38
Body part	6
Geography	14
Kinship	5
Nature	41

Object	13
Person/Age	6
Quality	56
Religion	32
Time	2
War/Politics	47
Unclassified	23

Table 5: Semantic categories of name elements and their distribution.

2.3. Fick's view[9]

> Die Masse der griechischen "Kurznamen" hat "Vollnamen" neben sich, und hierin liegt schon der Beweis, dass der einstämmige Name durch kosende Kürzung aus dem zweistämmigen entstanden ist. (Fick 1894: 36)

One could paraphrase Fick's claim as follows:

> 'The bulk of Greek monothematic names have dithematic names beside them' (Fick 1894: 36) *must mean that in the vast majority of cases, the element used for the formation of the monothematic name is derived from one of the elements of a dithematic name. If so, there should be a strong overlap of elements, even allowing for a certain rate of name loss.*

Three different complementary analyses of the data are used to examine Fick's claim with respect to the Elean corpus:

(1) examination of the distribution of elements,
(2) grammatical and semantic analysis of the elements,
(3) semantic and distributional analysis of the monothematic names formed with elements found in dithematic names.

---

[9] The discussions of both Fick's and Pulgram's view will be preceded by quotations from their work and my paraphrase (in *italics*) of the expectations arising from their point of view.

2.3.1. Distribution of elements

The first step is to extract the elements which monothematic and dithematic names share, as that is the common ground Fick must refer to.

Of the 289 different elements, 181 occur in monothematic names, 170 in dithematic names. Of these, 90 different stems occur in the first part of the names, and 108 in the second. 28 elements therefore can occur in either part of a compound name. 62 elements can occur in either a monothematic or a dithematic name. If we put together the elements occurring in monothematic and in dithematic names, the figures read as follows:

Elements' distribution	Number of elements	Percentage of the 181 elements used in monothematic names	Percentage of the 170 elements used in dithematic names
Only monothematic names	119	65.7	0
Monothematic and dithematic names	62	34.3	36.5
Monothematic names and both parts of dithematic names	21	11.6	12.4
Monothematic names and the first part of dithematic names	45	24.9	26.5, i.e. 50% of the first elements in dithematic names
Monothematic names and the second part of dithematic names	37	20.4	21.8, i.e. 34.3% of second elements in dithematic names

Table 6: Distribution of elements in monothematic and dithematic names

The most remarkable feature is thus that of the 181 elements used in monothematic names, about two thirds do not occur at all in any part of the dithematic names. In the remaining third, those elements which monothematic and dithematic names do share, it is noticeable that while 50% of the first elements in dithematic names recur in monothematic names, the same is true for only 34.3% of the second elements. This observation fits in with the long-standing hypothesis that the second element is more likely to be lost in the shortening process (cf. e.g. Locker 1933: 145, Masson 1995: 708, Schmitt

1995*a*: 425).[10] Taken together, the overlap in the Elean corpus does not support Fick's claim, as only a third of the elements used to form monothematic names in Elis have counterparts in dithematic names.

2.3.2. Grammatical and semantic analysis of the elements

Even though the overlap of elements in the Elean corpus does not support Fick's claim, one could defend his argument: since we have both monothematic and dithematic names from the earliest syllabic (Linear B) and alphabetic (Homer) evidence onwards, the claim must be pre-historic. One could therefore argue that the lack of overlap is due to the loss of names before the historic period.

In that case, a high consistency in the grammatical and semantic type of the monothematic and the dithematic names is to be expected, as the tendencies of name formation may be expected to remain the same, even if single elements get lost. In the following, we will therefore examine the distribution of the grammatical and semantic categories described above.

Taking the grammatical categories *Adjective, Proper Name, Substantive,* and *Verb*, which together account for 92.7% of the elements, we find the following distribution (Table 7):

Grammatical category	Percentage of elements in monothematic names	Percentage of elements in dithematic names
Adjectives	19.4	15.9
Proper Names	16.0	5.9
Substantives	54.7	58.2
Verbs	10.5	17.6

Table 7: Proportional distribution of grammatical categories

While the proportions of adjectives and substantives differ only insignificantly in monothematic and dithematic names, the same cannot be said for proper names and verbs, where the difference is 10.1 and 7.1 percentage points respectively.

In the semantic categories, similar observations can be made. Again, only the most common categories are taken into consideration:

---

[10] The relation remains remarkably unchanged when only those elements are taken into account which may occur in either the first or the second part of a dithematic name, but not in both: 34.8% to 18.4%.

Semantic category	Percentage of elements in monothematic names	Percentage of elements in dithematic names
Action	7.7	17.6
Nature	18.2	8.8
Quality	22.7	17.6
Religion	14.4	8.8
War/Politics	17.7	21.2

Table 8: Proportional distribution of semantic categories

Here, *Action* and *Nature* show the greatest differences with 9.9 and 9.6 percentage points divergence respectively. The divergences in distribution in both grammatical and semantic categories suggest that it is worthwhile to examine the distribution of the elements more closely, using not percentages, which as such do not permit conclusions (cf. Muller 1973: 52), but the *absolute discrepancy*. Rather than calculating only the distribution, as percentages do, with this method, it is first calculated what the distribution of elements in a sub-corpus should be if it were absolutely regular with regard to the corpus. This hypothetical distribution is then put against the observed distribution. The *absolute discrepancy* is then obtained by subtracting the hypothetical distribution from the observed one.[11]

In this way, we obtain absolute values which in our case can be used to estimate both the general tendencies in name formation for monothematic and dithematic names respectively, and the probability of derivation in the case of monothematic names.

2.3.2.1. Grammatical analysis: a closer look

The results of the calculations for expected and observed frequency and the absolute discrepancy are given in table 9 for the grammatical categories only:

---

[11] The formula (cf. Muller 1973: 53) for the expected value is the total of elements in the category multiplied by the total of elements in the sub-corpus. This number, divided by the total of elements in the corpus, provides the expected value E. For the verbs in purely monothematic elements, that would be (36 x 119) : 289 = 14.8, rounded up to 15. The 'absolute discrepancy' is then achieved by subtracting the expected value (E) from the observed value (O). In our example that is 6 - 15 = -9.

Grammatical category	Number of elements	Monothematic (119)			Dithematic (108)			Shared (62)		
		E	O	D	E	O	D	E	O	D
Adjective	48	20	21	**1**	18	13	**-5**	10	14	**4**
Proper Name	33	14	23	**9**	12	4	**-8**	7	6	**-1**
Substantive	151	62	62	**0**	56	62	**6**	32	27	**-5**
Verb	36	15	6	**-9**	13	17	**4**	8	13	**5**
Other	21	8	7	**-1**	9	12	**3**	5	2	**-3**

Table 9: Expected and observed frequency in the grammatical analysis of the elements
(E = expected frequency; O = observed frequency; D = absolute discrepancy)

Before we move on to the interpretation of these values, it seems advisable to establish the significance of the divergence through the χ^2 test.[12] χ^2 is obtained by the following calculation: Σ [(O-E)^2 / E]. The calculated χ^2 is then compared to the critical χ^2 in a χ^2 value table[13] under the correct number of degrees of freedom (df). The df are obtained through (number of rows – 1) x (number of columns – 1).

In this case, that is (5 – 1) x (3 – 1) = 8. The comparison is usually made at the 5 per cent level (an arbitrary value). If the calculated value is higher than the critical value, then the divergence is considered to be significant and the distribution cannot be considered to be due to chance.

For our data, χ^2 = 28.37 – the critical χ^2 for the 5% level and 8 df is 15.51, so that there is a significant difference in distribution even at the 0.1% level, where the critical χ^2 is 26.12.

Having established the significance, our interest now lies in the interpretation of the absolute divergences. The largest divergence is in the category *Proper Name*: it shows a substantial positive divergence in the monothematic elements, and a similarly high negative divergence in the dithematic elements. This implies that monothematic names were more commonly formed on the base of proper names, while dithematic names seem to avoid that type of formation. This indicates that the presence of a shared element does not in itself give sufficient ground to suppose that the monothematic name has to be an abbreviation of a dithematic name. Thus e.g. Ἀπολλώνιος and Ἀπολλόδωρος

---

[12] For more detail on the χ^2 test see e.g. Muller (1973: 116ff.), Butler (1985: 113ff.).
[13] Such tables can be found in any book on statistical linguistics, e.g. Muller (1973: 179), Butler (1985: 176), Woods-Fletcher-Hughes (1986: 301), as well as in the Internet under the term 'chi-square distribution'.

are likely to be independent formations built on Ἀπόλλων, a more plausible assumption also from a morphological point of view.

Another marked divergence is in the *Verbs*, which are far less common than expected in the monothematic elements, while more common in dithematic names. The positive divergence in the shared elements, which contrasts with the negative one in the monothematic elements, makes it quite likely that in the case of verbal elements in monothematic names, these may indeed be derivatives from dithematic names, where that category is common. One could even consider for those few monothematic names with verbal elements which lack a counterpart in the dithematic names, that the latter did get lost. In that case, a verbal element would be an indication for a derivation.

The negative divergence of -5 in *Adjectives* in the dithematic elements may have a semantic reason: adjectives with negative connotations such as 'clump-footed' are much less likely to occur in dithematic names (see above, section 1.1., on the 'ehrenvolle' connotation of elements). As there is a positive divergence in the *Adjectives* of the shared elements, these might well be derived, as may, but need not be, the *Substantives* in that category.

2.3.2.2. Semantic analysis: a closer look

As the values for the *Proper Names* in the grammatical analysis indicate, shared property alone is no guarantee of derivation of the monothematic names from dithematic ones. On the contrary, while derivation appears to be very likely for some categories, namely *War/Politics*, *Action*, and possibly *Quality*, it seems improbable for others, especially *Nature* and *Religion*.

Table 10 lists all 62 shared elements with the category they are assigned to:

Element	Semantic category	Element	Semantic category	Element	Semantic category
ἀγαθο-	Quality	θρασυ-	Quality	ὀνησι-	War/Politics
ἀγε-	War/Politics	ἱερο-	Religion	ὀνομο-	War/Politics
αἰχμο-	War/Politics	ἱππο-	Nature	Παιανο-	Religion
ἀλεξι-	War/Politics	καλλι-	Quality	πατρο-	Kinship
ἀλκα-	Quality	κλεινε-	War/Politics	πιστο-	Quality
ἀναξ-	War/Politics	κλειτο-	War/Politics	πολυ-	Quality
ἀνδρα-	Person/Age	κλεο-	War/Politics	Ποσειδο-	Religion
ἀνθο-	Nature	κρατε-	War/Politics	πρωτο-	Quality
ἀντι-	Prefix	κτητο-	Quality	Πυθο-	Religion
ἁπλο-	Quality	κυλλο-	Quality	πυρι-	Nature
Ἀπολλο-	Religion	λαμπο-	Quality	σαϝι-	War/Politics

ἀριστο-	Quality	λαο-	War/Politics	στρατο-	War/Politics
ἀρκεσι-	Action	λεοντο-	Nature	τεισι-	War/Politics
ἀρχε-	War/Politics	λυκο-	Nature	τελε-	War/Politics
Ἀφροδιτ-	Religion	λυσι-	Action	τιμη-	War/Politics
δημο-	War/Politics	μνα-	Action	τροφο-	Kinship
δρομο-	Action	μοσχο-	Nature	τυχε-	Quality
ζω-	Action	μουσο-	Religion	φιλο-	Quality
ἡγε-	War/Politics	νικο-	War/Politics	χαιρε-	Action
Ἡρα-	Religion	ξενο-	War/Politics	χαρι-	Quality
θηρα-	Nature	Ὀλυμπιο-	Religion		

Table 10: The shared elements and their semantic categories

More than 90% of the shared elements and 74% of all 289 elements belong to one of the five most common categories: *Action*, *Nature*, *Quality*, *Religion*, and *War/Politics*. Of these, the categories *Action*, *Quality* and *War/Politics* are most likely to contain elements complying with the description of dithematic names (e.g. Solmsen, von Kamptz, Masson, Schmitt, see sections 1.1. and 1.3. above), and therefore the names formed from those elements are more likely to be derived than those formed from elements of the other two categories. Elements from *Religion* in particular should not be included in the count as it is far more likely that the monothematic names were derived directly from the deity who in most cases is the base for the name, and not from a dithematic name making use of the same source (cf. section 2.3.2.1. above for *Proper Names*).

The comparison of expected and real frequency of the elements given in Table 11 supports these views (for background and formula see Table 9).

Semantic category	Number of elements	Elements in monothematic names (119)			Elements in dithematic names (108)			Shared elements (62)		
		E	O	D	E	O	D	E	O	D
Action	38	16	8	**-8**	14	24	**10**	8	6	**-2**
Nature	41	17	26	**9**	15	8	**-7**	9	7	**-2**
Quality	56	23	26	**3**	21	15	**-6**	12	15	**3**
Religion	32	13	17	**4**	12	6	**-6**	7	9	**2**
War/Politics	47	19	11	**-8**	18	15	**-3**	10	21	**11**
Other	75	31	31	**0**	28	40	**12**	16	4	**-12**

Table 11: Expected and observed frequency of elements in semantic categories

$\chi^2 = 56.9$ – the critical χ^2 for the 5% level and 10 df is 18.31, so that here again, there is a significant difference in distribution even at the 0.1% level, where the critical χ^2 is 29.59.

Let us now move to the interpretation of the *absolute discrepancies*, which are of greater interest here. In the elements used only in monothematic names, there are marked negative divergences in the categories *War/Politics* and *Action*. On the other hand, *War/Politics* has a high positive divergence in the shared elements. This indicates that those monothematic names containing elements from the category *War/Politics* could indeed be derivatives from dithematic names: they do not appear to be normally used for the formation of monothematic elements. Therefore, where they do occur, a derivation is probable.

Action shows a similar pattern for elements used only in monothematic or dithematic names respectively, but a slight negative divergence in the shared elements. *Action* elements seem therefore to have been avoided in any monothematic name formation, which tallies with *Verbs* being avoided in the grammatical categories (though the two are not always identical).

The positive divergence in elements used solely in monothematic names in the category *Nature* indicates that elements from this semantic category appear to be more likely to create originally monothematic names than derived monothematic names.

The categories *Quality* and *Religion* show similar divergences to one another. The negative discrepancy in the dithematic names indicates that both categories are less likely to occur in elements purely used in dithematic names. This tallies both with the expectation that certain qualities such as πιστός are common in monothematic slave or nick names (cf. e.g. Lambertz 1908: 22, Bechtel 1917: 503) and with the observation concerning *Proper Names* in the grammatical categories.

2.3.3. Names vs. elements

Fick is not talking of elements, but of names. It is therefore necessary to examine the names formed from the elements shared by mono- and dithematic names.

Here, the statistics are somewhat different: 130 or 47.1% of the 276 Elean monothematic names are formed using the 62 elements shared with the dithematic names.

This is due mainly to the fact that one element or root can be combined with different suffixes to form a number of different names: 10 of the 62 shared elements account for four or more of the monothematic names each, together

making up 43.1% of the 131 formed from the shared elements. Table 12 shows the distribution of shared elements over monothematic names:

Element	Number of names		ἀριστο-	4
φιλο-	15		δημο-	4
λυκο-	6		θρασυ-	4
ἀλεξι-	5		ἱππο-	4
καλλι-	5		λεοντο-	4
νικο-	5			

Table 12: The number of different names formed from the most common shared elements

However, here again we must take into consideration the different semantic categories of the elements from which the names are formed. Of these, ἀλεξι-, ἀριστο-, δημο-, καλλι-, νικο-, and φιλο- are all elements which express positive qualities or refer to politics, warfare, or heroic deeds. All of them are used even within the Elean corpus in typical compound names such as Ἀλέξανδρος, Ἀριστόμαχος, Δαμόνικος, Καλλίστρατος, Νικόμαχος, Φιλό-νικος, all names conforming to the Greek ideas of honour and glory. A genuine derivation of the monothematic names featuring these six elements from dithematic names is therefore possible, and perhaps even probable.

The same is true for the element ἱππο-: the horse played an important part in Greek war strategy and sports, being connected with the image of bravery and speed.[14] In the Elean names, it occurs in fourteen different dithematic names, among them e.g. Ἵππαρχος, Ἱππόμαχος and Νικάσιππος.

This contrasts sharply with the other two animal names common in monothematic Elean names, λεοντο- and λυκο-. λεοντο- is found in five dithematic names, and λυκο- only in one. It seems therefore very likely indeed that the monothematic names from these last two animals, comparable to the names formed with *Religion* elements, are constructed directly from the name of the animal, rather than via a dithematic name.

2.3.4. Conclusion

The analysis of the Elean corpus does not support Fick's theory: two thirds of the monothematic names show elements not occurring at all in dithematic names. Of the remaining third, which monothematic and dithematic names

---

[14] Cf. again the story of Pheidippides: 'She (Pheidippes' mother) wanted a *hippos* attached, a knight's name' (Ar. *Nub.* 60ff.).

share, the elements belonging to the grammatical category *Proper Name* or the semantic categories *Religion* and *Nature* (the latter with the possible exception of the element ἱππο-) are unlikely to have been used in monothematic names as a result of derivation. The number of monothematic names formed from the elements shared by monothematic and dithematic names[15] amounts to only 36.7% of the total of names, which is far from the 'bulk' to which Fick refers.

The Elean corpus, then, considered from a non-historical point of view (and, like most epigraphic corpora, necessarily incomplete), confirms two suspicions long held by scholars: Fick's theory of a near-complete derivation of the monothematic names from the dithematic ones is not supported by the facts, and presents a restricted view of the Greek naming system, not fully acknowledging its breadth and depth.

However, in those cases in which derivation is possible, it is more likely to have happened using the first element of compound names rather than the second.

2.4. Pulgram's view

> In accordance with universal traits of nomenclature and in view of the serious gaps in the evidence we possess, we must suppose that the oldest form in Indo-European was a simple appellative and not a compound (Pulgram 1947: 206)

With reference to Greek in particular, Pulgram says:

> At that point [i.e. during the second millennium BC], the Greek-speaking lords of the feudal society (which is without doubt characteristic of Greece in the Bronze Age, and is clearly described in the Homeric poems) could be sufficiently distinguished from the non-Greeks by their Greek names alone, without giving them a special shape. Only when every Tom, Dick, and Harry of Hellas, whether or not a descendant of the original Hellenizers of the land, could sport a Greek name, those who thought of themselves as aristocrats shifted to another naming device which would again separate them from the common crowd. At that time, they acquired the dithematic names which, some centuries later again, are borne by most Greeks' (Pulgram 1960: 201)

The reasons for his refutation of Fick's theory, therefore, lie mostly in an *argumentum ex silentio* on social grounds:

---

[15] Excluding the elements from 'Religion' and 'Nature' (but including 'ἱππο-') from the count of names for the reasons named above in 2.3.3.

[There is] an inequitable representation of the common people in extant documents. (Pulgram 1947: 205)

[...] the numerical predominance of dithematic names [...] is statistically misleading because it fails to take into consideration the frequency-ratio of each individual name. [...] (Pulgram 1947: 205)

That the formal difference between monothematic and dithematic names reflects also one of social character can be proved for Greek by the fact that monothematic names were borne, with particular frequency, by male and female slaves and by harlots, and that free persons were named so only late, from the time of the monarchy on (Pulgram 1947: 202)

The prevalence of compounds in available records for some areas is not necessarily typical for any speech-area as a whole, and even less for an entire linguistic family, in view of the socially and historically conditioned partiality of the tradition which favours the upper classes. (Pulgram 1947: 206)

The implications of Pulgram's statements appear to be the following:

There is a sharp social division concerning the use of monothematic and dithematic names in Greek, the former being used for 'common' *people, the latter in aristocratic circles. As documentation favours the upper classes, dithematic names are expected to make up the bulk of the evidence. Monothematic names are expected to occur with much greater frequency* 'from the time of the monarchy on' (Pulgram 1947: 202)*, when the higher classes first started to use that type of naming again. The frequency-ratio for individual names should be checked.*

The Elean material does not support these social arguments, as becomes apparent from the examination of the corpus from the following points of view:

(1) Numerical relation of mono- and dithematic names,
(2) Distribution of mono- and dithematic names over the period of evidence (800 BC - AD 400),
(3) Frequency-ratio of individual names.

2.4.1. Numerical relation of mono- and dithematic names

As we have seen above, the relation of different monothematic to dithematic names is 276 to 312, i.e. 46.9% to 53.1%. Although favouring dithematic names

slightly, this can hardly be interpreted as showing the 'prevalence of compounds' mentioned by Pulgram (1947: 206).

2.4.2. Distribution of mono- and dithematic names over time

The 588 different Greek names are borne by 912 individuals, with 26 occurrences of an individual bearing more than one name, raising the number to 938 names shared. These are distributed over a period of roughly 1250 years. Following the practices of archaeology and ancient history, we may distinguish five major periods (Table 13):

Time	Name of Period
800 - 500 BC	Archaic Period
500 - 333 BC	Classical Period
333 - 146 BC	Hellenistic Period
146 - 32 BC	Roman province in Late Republic
32 BC - AD 284	Early Roman Empire (Principate)
AD 284 -	Later Roman Empire (Dominate)

Table 13: The periods in the Elean corpus

The distribution of names in these periods is as follows (Table 14):

Period	Greek names (938)	Monothematic names (445)	Dithematic names (493)
Archaic Period	14	13	1
Classical Period	107	49	58
Hellenistic Period	209	104	105
Roman province in Late Republic	93	38	55
Early Roman Empire (Principate)	482	226	256
Later Roman Empire (Dominate)	7	1	6
Undatable	22	12	10

Table 14: The distribution of names over time

The period with by far the most material is the Principate, accounting for 51.6% of the corpus. It is followed by the Hellenistic Period (22.4%). The Classical Period accounts for 11.5%, the Late Republic for 10%. There is very little material from the earliest and latest times: the Archaic Period and the Dominate *together* amount to only 2.2%. Similarly, the undatable material accounts for 2.3%.

The distribution of monothematic and dithematic names differs markedly from the one we would expect taking into account Pulgram's arguments above: if monothematic names were used predominantly in the lower classes, but the upper classes with their dithematic names dominate the extant documents, we would expect a rather low percentage of monothematic names. However, the four periods with the largest amount of data, namely the Classical Period to Early Roman Empire, show high percentages of monothematic names throughout: Classical Period 45.8%, Hellenistic Period 49.8%, Roman province in Late Republic 40.9%, and Early Roman Empire 46.9%.[16]

This suggests either that the difference between monothematic and dithematic names was socially less distinctive than Pulgram suggests,[17] or else that 'common' people are mentioned in the documents more frequently than previously assumed. The fact that there is no sharp increase in monothematic names in the later periods — as we would expect if the development were what Pulgram believed it to be — may also indicate a less pronounced social division in the names.

This is another example which demonstrates that although names can often serve as a useful indicator of social status, the social implications may be more complicated than they at first appear: in Germany, for example, it is currently fashionable to give exotic or rare names in order to distinguish the child from birth onwards. The trend is probably more pronounced among celebrities and, in imitation, in the lower classes. Names recently approved by the authorities included *Nica-Luna*, *Seraina*, and *Navik* (cf. 'Unser Kind heisst Maise' 2001). Among those not approved were e.g. *Borussia*, *Uragamo*, and *Scheisserle* ('little shit') (cf. Rupp 2000). This practice reminds one of literary examples of motives for naming: 'Her real name is Marilyn Florence Isabel. Mrs. Snowbeam says the only thing she could give her children was real fancy names.' (Montgomery 1937: 101).

---

[16] Note that the high percentage of monothematic names in the Archaic Period (92.2%) should be treated with extreme caution for two reasons. Firstly, because the names of that period only make up 1.5% of the corpus. Secondly, because only 4 of the 13 monothematic names, i.e. just under a third, are attested in inscriptions, while all others derive from the victors' lists, compiled much later than the dates they refer to.

[17] Cf. Bechtel (1898: 6) about Πλάτων (cf. note 3).

2.4.3. Frequency-ratio of individual names

Pulgram claims that the statistical data concerning monothematic and dithematic names are misleading as long as they do not take into account the frequency of the individual names. In the Elean corpus, the frequency-ratio is as follows (Table 15):

Borne by	Monothematic Names (276)	Dithematic Names (312)
more than 10	2	0
8	1	2
7	1	2
6	3	5
5	4	3
4	5	8
3	27	19
2	35	56
1	198	217
Frequency-ratio of names of individuals (938)	47.4%	52.6%
Frequency-ratio of *different* names (588)	46.9%	53.1%

Table 15: Frequency ratio of individual names

Table 15 indicates that on the whole names were less repetitive than we commonly assume: only two names in the whole corpus are borne by 10 individuals or more: Διονύσιος (13 times) and Ἀπολλώνιος (10 times). Interestingly, they are both monothematic, theophoric names, and, with one exception, occur only in the Principate. The one possible exception is the first example of Ἀπολλώνιος, which appears to date from the Late Republic.[18]

The following table analyses the actual frequency-ratio of mono- and dithematic names against the expected ratio.

---

[18] Διονύσιος heads the most popular names in all published volumes of the *LGPN* (I-IIIB). In all these volumes, Ἀπολλώνιος is among the most popular names, in vols. I-IIIA even among the first 10. Their popularity in Elis might be an indicator of the representative nature of the Elean corpus in a wider Greek context.

Number of bearers	Number of names in category	Monothematic (445)			Dithematic (493)		
		E	O	D	E	O	D
4 or more	36	17	16	-1	19	20	1
3	46	22	27	5	24	19	-5
2	91	43	35	-8	48	56	8
1	415	197	198	1	218	217	-1

Table 16: Expected and observed frequency of names

The most marked difference lies in the names occurring twice, of which there are fewer monothematic and more dithematic ones than expected. On the whole, however, there is no indication that monothematic names in general showed a higher frequency than dithematic ones, and the variety of names in the Elean corpus appears to reflect the wealth of Greek onomastics.[19]

2.4.4. Conclusion

From the above, it becomes obvious that Pulgram's *argumentum ex silentio* from social implications in favour of monothematic names is not supported by the Elean corpus. While his refusal of Fick's theory agrees with the Elean data, his reversal of Fick's claim does not.

Indeed, one could also suppose that common people would have given their children grand, dithematic names, in order to indicate their children's individuality, or to side with Mrs. Snowbeam and at least give their offspring 'real fancy names', if they could not offer them much else. This would actually tally with the account of the naming of *Pheidippide*s, whose father wanted to remind him in his name of his origins and background, while his mother wanted a dithematic name with noble associations.[20]

---

[19] Interesting from a different point of view is the fact that nearly 80 per cent (78.4%) of the names occur only between one to three times. This may be an indication that possibly *Nachbenennung* as a motive for naming (e.g. calling one's son after one's father) was less common than we assume, as even with a high loss of material we might expect a higher frequency-ratio for names in general. It might be worth examining if instead the use of elements from the parents' or grandparents' names was more widespread.

[20] One could even take this fantasy further and come to a conclusion diametrically opposed to Pulgram's: if we assumed that the 'common people' gave their offspring dithematic names for the reasons referred to above (2.4.2.), but still could not pay for any sort of memorial, then all those missing names Pulgram talks about would actually have to be dithematic! Non-existent

3. Final observations

The paper has two aims: firstly, the definition of a non–historical approach to Greek names which avoids earlier, often ambiguous nomenclature, and allows a more objective categorization. Secondly, the statistical analysis of the grammatical and semantic categories of a temporally and geographically defined corpus which tests, and ultimately speaks against, the extreme theories of Fick and Pulgram as regards the relationship of monothematic and dithematic names. The analysis substantiates the view long held by Solmsen, Masson, von Kamptz, and Schmitt, that there are both originally monothematic and originally dithematic names, and in addition to these, derived monothematic names.

This threefold system of monothematic, dithematic and derived monothematic names, which is supported by the analysis of the Elean material, appears to be very common, and it is admirably understood and expressed by the Gaffer in the last chapter of the *Lord of the Rings*, when giving advice on the naming of his granddaughter: 'Make it short, and then you won't have to cut it short before you can use it!'

References

Bechtel, Friedrich (1898). *Die einstämmigen männlichen Personennamen des Griechischen, die aus Spitznamen hervorgegangen sind*. Berlin.

— (1917). *Die historischen Personennamen des Griechischen bis zur Kaiserzeit*. Halle.

Butler, Christopher (1985). *Statistics in Linguistics*. Oxford.

Davies, J.K. (1971). *Athenian Propertied Families: 600-300 BC*. Oxford.

Dobias, Catherine, and Dubois, Laurent (1990). 'Introduction', in Olivier Masson, *Onomastica Graeca Selecta I*. Nanterre, I-XVI.

Fick, August (1894). *Die Griechischen Personennamen nach ihrer Bildung erklärt und systematisch geordnet* (2nd edn. by Friedrich Bechtel and August Fick). Göttingen.

Fraenkel, Ernst (1935). 'Namenwesen. Die griechischen Personennamen'. *RE*, 16: 1611-48.

Kamptz, Hans von (1982). *Homerische Personennamen. Sprachwissenschaftliche und historische Klassifikation*. Göttingen.

Lambertz, Max (1907/8). *Die griechischen Sklavennamen*. Wien.

---

material, and conclusions drawn from it, should therefore only be taken into consideration with the utmost care.

LGPN = Fraser, Peter, and Matthews, Elaine (1987-). *A Lexicon of Greek Personal Names.* Oxford.

Locker, Ernst (1933). 'Die Bildung der griechischen Kurz- und Kosenamen'. *Glotta*, 21: 136-52.

— (1934). 'Die Bildung der griechischen Kurz- und Kosenamen'. *Glotta*, 22: 46-100.

LSJ = Liddell, Henry George, Scott, Robert, and Jones, Henry Stuart (1996). *A Greek-English Lexicon* (9th edn.). Oxford.

Masson, Olivier (1990, 2001). *Onomastica Graeca Selecta* (3 vols., ed. C. Dobias and L. Dubois). Nanterre.

— (1995). 'Les noms propres d'homme en grec ancien', in *NF*, vol. 1, no. 102: 706-10.

Meister, Richard (1890). 'Vollname und kurzname bei denselben personen überliefert', *Beitr. Idg. Spr.*, 16: 173-4.

Montgomery, Lucy Maud (1989 [1937]). *Jane of Lantern Hill.* New York/Toronto.

Muller, Charles (1973). *Initiation aux méthodes de la statistique linguistique.* Paris.

NF = Namenforschung/Name Studies/Les noms propres. Ein internationales Handbuch zur Onomastik/An International Handbook of Onomastics/Manuel international d'onomastique (2 vols., ed. Ernst Eichler, Gerold Hilty, Heinrich Löffler, et al.). Berlin/New York 1995.

Notopoulos, James A. (1939). 'The name of Plato'. *Classical Philology*, 34: 135-45.

Pulgram, Ernst (1947). 'Indo-European personal names'. *Language*, 23: 189-206.

— (1960). 'New evidence on Indo-European names'. *Language*, 36: 198-202.

Rupp, Hugo (2000). 'Mein Name ist Schulze, Kimberley Schulze'. *Die Zeit*, 20/2000.

Schmitt, Rüdiger (1995*a*). 'Morphologie der Namen: Vollnamen und Kurznamen bzw. Kosenamen im Indogermanischen', in *NF*, vol.1, no. 61: 419-27.

— (1995*b*). 'Entwicklung der Namen in älteren indogermanischen Sprachen', in *NF*, vol. 1, no. 94: 616-36.

Solmsen, Felix (1922). *Indogermanische Eigennamen als Spiegel der Kulturgeschichte* (ed. and rev. by E. Fraenkel). Heidelberg.

Tolkien, J. R. R. (1955). *The Lord of the Rings*. London.

'Unser Kind heisst Maise' (2001). *Focus*, 45/2001: 231-2.

Wood, Anthony, Fletcher, Paul, and Hughes, Arthur (1986). *Statistics in Language Studies*. Cambridge.

Phonological reconstruction and the rôle of semantics in etymology

The case of Greek ἥκω and ἵκω 'to (have) come'

Andreas Willi

1. Introduction

The development of sound methodological principles in language comparison and reconstruction is perhaps the greatest achievement of nineteenth-century linguistics. Modern comparative philology is based on the recognition of regular sound changes, and the neogrammarians' concept of 'sound laws' is the backbone of the entire discipline. The nineteenth-century focus on phonological comparison was a healthy reaction to earlier forms of linguistic comparison which were mostly based on the lexicon and therefore semantic by nature.[1] However, the methodological progress had an ambivalent impact on comparative research until this day.

The more sound laws were recognized the more peripheral became the role of semantics. This had two reasons. First, the regular application of sound laws to two semantically disagreeing forms could prove that they had a common source. In such cases, the only semantic task was in a second step to find a plausible chain of semantic developments. An example is the relationship between Greek δόρυ 'spear' and German *treu* 'faithful': the semantic link lies in the fact that the Greek spear could be made from oak wood and that oak wood is particularly hard or reliable.

The second reason for the diminished rôle of semantics was more serious. Two semantically agreeing *and* phonetically similar words could be shown to be genetically *un*related. The most famous case is that of Greek θεός 'god' and Latin *deus* 'god' (Morpurgo Davies 1998: 172–3). Their semantic equivalence is so straightforward that August Pott, in 1833, felt obliged to postulate, beside the well-attested phonetic equivalence of Greek δ- with Latin *d*- (cf. Gr. δόμος = Lat. *domus* 'house'), a second pair with Greek θ- corresponding to Latin *d*-: that is, he did what nowadays would amount to the reconstruction of a further Proto-

---

[1] Even though the question of *litterarum permutationes* or phonetic equations had always played some role since it would have made little sense to compare two words *only* because of their common meaning (cf. Morpurgo Davies 1998: 47-9).

Indo-European phoneme. It took almost 30 years until Georg Curtius dared to reject the connection. Here, then, semantics was not only relegated to the second rank but actually shown to be unreliable and misleading.

The example of θεός and *deus* shows that the consideration or non-consideration of semantic similarity can have a crucial impact on phonological reconstruction. It is clear that this was more of a burning issue when the phonological system of Proto-Indo-European was less well explored than it is today. Yet, there are still situations in which we end up with fundamentally different views on the Proto-Indo-European sound system depending on 'how much semantics we take with our phonology'.

In the present paper I will illustrate this continuing dilemma with an old problem of Greek historical linguistics. First, I will review two competing solutions, one of which is formally more rigid and therefore 'anti-semantic', while the other allows for more formal flexibility (not to say fuzziness) in order to remain 'pro-semantic'. The case is exemplary because it also reflects a part of the history of Indo-European studies. The 'pro-semantic' solution is advocated by French-speaking linguists who tend to be more interested in Meillet's 'histoire des mots'. The 'anti-semantic' solution is supported by German or German-speaking scholars who may still be more thoroughly impregnated with the neogrammarian spirit. After a short critical discussion I will present a new hypothesis, which stands somewhere in between the two and tries to be 'pro-semantic' and 'pro-formal' at the same time.

2. *Greek ἵκω, ἱκνέομαι, ἱκόμην and Proto-Indo-European* *seik-/*sik-

The problem to be dealt with is the relationship between the two Greek presents ἥκω and ἵκω. Both ἥκω and ἵκω mean not just simply 'to come' but also, with some kind of a perfective nuance, 'to have come, to have reached'. For instance, when Odysseus meets Athena in the shape of a young shepherd after his arrival in Ithaca, he asks her where he is and the goddess replies as if she were surprised at the question since 'the name of Ithaca has even reached the land of Troy' (*Od.* 13.248: Ἰθάκης γε καὶ ἐς Τροίην ὄνομ᾽ ἵκει). It is clear that Athena does not want to say that the name is currently reaching Troy (which is destroyed at that point).[2]

Ἥκω is the regular form in Attic Greek, whereas in Homer and in Doric Greek ἵκω is more common. The anlaut of ἵκω is long, but there are two

---

[2] For similar examples with ἥκω see *A Greek-English Lexicon*, compiled by Henry George Liddell and Robert Scott, revised and augmented throughout by Sir Henry Stuart Jones (9th edn.), Oxford 1940, 767, s.v. ἥκω (translated with 'to have come, be present').

apparently related nasal stems ἱκνέομαι and ἱκάνω with a short ἱ-. Similarly the unaugmented forms of the thematic aorist ἱκόμην show a short anlaut ἱ-. All of these, ἱκνέομαι, ἱκάνω, and the aorist ἱκόμην, mean 'to arrive at, to reach'. Note that they do *not* share the perfective nuance of the presents ἥκω and ἵκω.

If we reconstruct a root **seik-/*sik-*, it is possible to connect the long-vocalic present ἵκω and the short-vocalic ἱκνέομαι and ἱκόμην. The present ἵκω would reflect a reduplicated stem **si-sik-*, while the other forms present the simple root (Klingenschmitt 1975*a*: 75). The resulting system can be compared with that of the root **segh-* in ἔχω:

ἵκω < **si-sik-*	ἱκνέομαι < **sik-n-*	ἱκέσθαι < **sik-*
ἴσχω < **si-sgh-*	-ισχνέομαι < **-si-sgh-n-*	σχεῖν < **sgh-*

The only difference lies in the nasal stem, but there the reduplicated -ισχνέομαι is more exceptional than ἱκνέομαι, and the latter has a parallel in another Indo-European language: Tocharian B *sikna-* 'to set foot, to step' shows, like Greek ἱκνέομαι, a nasal stem. Apart from this, the root **seik-/*sik-* also appears in Umbrian (*pru-sikurent* 'to confirm solemnly') and Lithuanian (*síekti* 'to reach out for sth.').[3]

The basic meaning of **seik-/*sik-* seems to be 'to reach out with either hand or foot', i.e. 'to reach, to grasp' or 'to step (onto/into)'. In Greek, the former option may be retained in the specialized meaning of ἱκνέομαι 'to be a suppliant': here ἱκνέομαι can have a direct object, and it is plausible that the phrase τὰ σὰ γοῦνα ἱκόμεθα in the *Odyssey* (9.267) literally means 'we reached out with our hands for your knees' (in the gesture of supplication: cf. Benveniste 1969: 252–4, Létoublon 1985: 151). The latter meaning 'to step' is seen in the noun ἴχνος 'track, footstep, trace', which has been explained as representing **sik-sn-* (Panagl 1976; cf. e.g. λύχνος 'lamp' < **luk-sn-*, Lejeune 1972: 74). A subsequent semantic shift from 'to step (upon sth.)' to 'to arrive at, to reach' is reasonably straightforward.

However, there are a few difficulties with the explanation of ἵκω via **seik-/*sik-*. First of all, it is surprising that a reduplicated present ἵκω should have a perfective meaning when the pure root, as represented in the nasal present ἱκνέομαι and the aorist ἱκόμην, has not.

Second, in Homer the long anlaut of ἵκω always stands *in arsi*. That is, the Homeric poems preserve no trace of an uncontracted intermediate form **hi-hik-*.

---

[3] Fraenkel (1962–5: 781, s.v. *síekti*), Hamp (1982: 70–2), Meiser (1986: 87–8), Rix et al. (2001: 522, s.v. **seik̑-*), Adams (1999: 691, s.v. *sik-*). Greek ἐνεῖκαι 'to bring' (Ionic aor. of φέρω) must probably be separated from this root (Meier-Brügger 1987, Strunk 1988: 576–80).

Of course there are other Homeric words where contraction of two similar vowels after the fall of an intervocalic *-s- is metrically guaranteed,[4] but ἵκω also occurs in the line-end formula (ἀϋτὴ δ᾽ εἰς/σέλας δ᾽ εἰς/κλέος) οὐρανὸν ἵκει/-ε(ν)/-η and such formulae tend to preserve archaic forms.[5] On the other hand, if the formula is recent (so that ἵκω with a long vowel in the anlaut could more easily show a relatively late contraction), it should belong to the Ionic layer of Homeric diction — even though, as Ruijgh (1957: 132–3, 1968: 119) has pointed out, an Ionic poet would be expected to use ἥκω in the first place, and ἵκω would make better sense if it belonged to the Aeolic or even Achaean layer of the poems.[6]

We will come back to this point later, but let us discard the objection for the moment. Let us also dismiss as a coincidence the fact that all the Homeric examples of the long-vocalic present stem ἱκ- show an active inflection, whereas all the examples of the short-vocalic aorist stem ἱκ- are in the middle voice. And let us finally overlook that ἵκω never takes a preverb, whereas ἱκόμην and ἱκάνω do so very frequently.[7]

3. Phonology vs. semantics: traditional explanations of Greek ἥκω

But how does ἥκω fit into the picture? We can reconstruct its root as *seh₁k-, which looks uncomfortably similar to *seik-, but in a conservative reconstruction of Proto-Indo-European phonology there is no possible way to derive one from the other.[8] This is the point where opinions diverge.

---

[4] Even in a similar environment, when verbs beginning with *se- are augmented (Chantraine 1958: 42).

[5] *Il.* 2.153; 2.458; 8.192; 8.509; 12.338; 14.60; *Od.* 9.20; 15.329; 17.565; *h. Apoll.* 442. For further details on the distribution of ἵκω/ἱκάνω etc. see R. Führer in the *Lexikon des frühgriechischen Epos*, 2.1167-70, s.v. ἱκάνω (where, somewhat inconveniently, ἵκω and ἱκάνω are treated together, while ἥκω is kept apart).

[6] The Ionic ἥκω is the vulgate reading only in *Il.* 5.478 and *Od.* 13.325.

[7] Cf. Létoublon (1985: 152) with further interesting observations: ἵκω occurs rarely, ἱκάνω often with an animate subject; with ἱκάνω inanimate subjects are usually palpable objects, with ἵκω impalpable ones (although it is also clear that the two verbs have influenced each other).

[8] Schmid (1956: 229) suggests that ἥκω might be a cognate of Skr. *yā́ti* 'to go' with a 'k-Erweiterung'. Since *yā́ti* probably goes back to *h₁i-eh₂- (cf. e.g. lat. *iānus*, Mayrhofer 1992–6: 2.407–8, s.v. YĀ ¹), this would mean that the West Greek and Aeolic attestations of ἥκω must be explained as due to Attic (or Koine) influence. While this is possible for most of them (e.g. *GDI* 2151 ποθήκω from Delphi; Theocr. 4.47), it seems difficult with Sappho fr. 114 L.-P. (cf. Ruijgh 1957: 132). Moreover, the reservations formulated in the main text against the reconstruction of *seh₁k- also apply, *mutatis mutandis*, to this variant of separating ἥκω and ἵκω from each other.

The new German *Lexikon der indogermanischen Verben* by Helmut Rix and his collaborators posits two unrelated roots **seik-* and **seh₁k-* (Rix et al. 2001: 519, s.v. *?*seh₁k-*, and 522, s.v. **seik̑-*). The only sign of hesitation is a question-mark in front of **seh₁k-*, probably because (1) the separation of ἥκω and ἵκω is semantically counterintuitive, and (2) there is no further compelling evidence for such a Proto-Indo-European root **seh₁k-*.[9]

On the 'French side', Pierre Chantraine (1968–80: 409, s.v. ἥκω) writes in his etymological dictionary of Greek: 'On pose **sēq-* ou **sēi(q)-* et on rapproche ἵκω, etc.'[10] This is extreme shorthand for a rather complicated theory. Chantraine probably implies the following: the **-k-* in the roots **seik-* and **seh₁k-* does not belong to the 'real' root but is a root enlargement. If we neglect it, we have an alternation **seh₁-/*sei-*. The latter, **sei-*, could also be reconstructed as **seHi-* with an intervocalic laryngeal. Thus, we end up with a vocalic alternation **seh₁-/*seh₁i-*, or, in non-laryngealist terms, with an alternation between a long vowel and a diphthong. Such an alternation seems to have parallels elsewhere (cf. Benveniste 1962: 147–73).

Let us take the root **peh₃-* 'to drink' (cf. Mayrhofer 1986: 173–5). We find the full grade for instance in Lat. *pōculum* 'cup' < **peh₃-tlom* and in the Greek imperative πῶθι < **peh₃-dhi* 'drink!'. Next to this, there is a second Greek imperative: πῖθι 'drink!'. In order to explain the latter we must start from a root variant **peh₃i-*, the zero grade of which provides an imperative form **ph₃i-dhi*. In **ph₃i-dhi* the so-called 'laryngeal metathesis' must have operated so that we finally reach **pih₃-dhi*, the regular ancestor of πῖθι. The explanation of words such as Greek ποιμήν 'shepherd' from a root **peh₂i-* (or **peh₃i-*?) beside Latin *pāstor* 'shepherd' from **peh₂-*, or of Greek θῆσθαι 'to suckle' from **dheh₁-* beside Luv. *titaimi* 'son, child' and Skr. *dhītá* 'suckled' from **dheh₁i-* must be sought along similar lines.

However, a closer examination shows that the assumption of an alternation between a long vowel and a diphthong has its disadvantages too. In laryngeal roots such as **peh₃i-* or **dheh₁i-* the 'additional' **-i-* occurs at the *end* of the root. The standard explanation for the **-i-* therefore sees its origin in the formation of present stems in **-i̯-* (Mayrhofer 1986: 174). A priori, the stem-formational suffix **-i̯-* should not then be added *before* a root enlargement in **-k-*. And even supposing that would work (since there may be the odd

---

[9] Lamberterie (1990: 294–5), with less hesitation than Klingenschmitt (1975*a*: 77 n. 5 and 1982: 213), suggests a derivation of Armen. *has-* 'to arrive' from **sh₁k-* (against the traditional **h₁nk-* from **h₁enk-*). This is now tentatively adopted by the second edition of the *Lexikon der indogermanischen Verben* (Rix et al. 2001: 519, s.v. *?*seh₁k-*), after García Ramón (1999: 58–9).

[10] Similarly Frisk (1960–72: 1.628, s.v. ἥκω ['idg. *sē(i)q-*?'], and 1.720, s.v. ἵκω).

parallel[11]), how would we reach a zero-grade *sik-?* If *seik-* is in reality *seh₁ik-*, its zero-grade should be *sh₁ik-*. Here the same laryngeal metathesis as in πῖθι would operate and ultimately yield *sih₁k-*, not *sik-*. What we need instead is a theory in which we have *either* the laryngeal *or* the *-i̯-*.

Another French scholar has proposed such a theory. Extending an idea about the laryngeal *-h₃-* by Martinet (1964: 212–34), Lamberterie (1990: 298–9) suggests that an intervocalic laryngeal *-h₁-* developed into an intervocalic *-i̯-* at a very early stage. Thus, a root *seh₁-* would have alternated between *seh₁-C-* and *sei̯-V-*. Subsequently (but still in Proto-Indo-European times) a split would have taken place and *seh₁-* and *sei̯-* would have been treated as separate roots.

Again, there are several objections to be raised. First, it is not particularly plausible (1) that, after such a split had taken place, both roots should have retained exactly the same meaning, (2) that the same root enlargement should have been added independently to two synonymous roots, (3) that they should still have remained synonymous even after this addition, and (4) that any language group (in our case: Greek) should have maintained the superfluous redundancy right down into the historical period.

Second, and more seriously, there is good evidence that intervocalic *-h₁-* did not develop into *-i̯-*: Vedic *vā́ta-* 'wind' from the root *h₂weh₁-* 'to blow' is often measured as a trisyllabic word (i.e. /vaHata-/) and must go back, like Latin *ventus* 'wind', to *h₂weh₁-n̥t-o-*; with an intervocalic change *-h₁-* > *-i̯-* we should expect *vayata-*, not *vā́ta-* (see Mayrhofer 1986: 124, with further points).

But even if these obstacles to Lamberterie's version did not exist, we might hesitate to reconstruct a phonological development of Proto-Indo-European on the basis of little more than the bizarre behaviour of one or two Greek verbs.[12] Do we therefore have to return to the 'German' anti-semantic position and separate ἥκω and ἵκω completely? Perhaps not.

4. Separating ἵκω and ἱκόμην

To start with semantics, we have seen that the perfective nuance of the presents ἥκω and ἵκω is surprising if it is compared with the non-perfective meaning of the root *seik-/*sik-* of ἱκάνω/ἱκνέομαι and ἱκόμην; semantically there is no

---

[11] Klingenschmitt (1982: 213 n. 69) refers for instance to Greek σκίπων 'staff' < *skih₂-p- beside σκῆπτρον etc. < *skeh₂-p-.

[12] Lamberterie himself has reservations about his 'reconstruction ultime' even though he also wants to solve with it the problem of Greek εὐθύς and ἰθύς vs. Skr. *sādhú-* (on which cf. Willi 2001).

doubt that ἱκόμην is the aorist of ἱκάνω/ἱκνέομαι, *not* of ἵκω.[13] We have also seen that the divergent attitudes towards composition with preverbs and the differences between the active and middle inflection of ἵκω and ἱκόμην respectively make the link even more problematic (cf. section (2) above).[14] The treacherous superficial similarity of ἵκω and ἱκόμην must not mislead us; both semantically and morphosyntactically ἥκω and ἵκω have more in common with each other than either of the two has with ἱκόμην and the like. Hence we should rather try to put together ἥκω and ἵκω while leaving aside ἱκνέομαι and ἱκόμην. *If* there is a formally unobjectionable way of relating ἥκω to ἵκω and finding an acceptable Indo-European pedigree for them, we should take it since it is ἥκω and ἵκω which are creating our etymological problems, not ἱκνέομαι with its well-established root *sik- 'to reach with the hand, to step'.

5. An old new root for ἥκω

Let us first look at ἥκω. There is no obligation to reconstruct its root as *seh₁k- — as long as there is any other root which would have yielded ἥκω by regular sound changes. After all, we do not want to invent a further *ad hoc* solution, for otherwise we might just as well accept either that of Chantraine or that of Lamberterie.
The initial aspiration of a Greek word need not go back to an initial *s-, it can also stem from *Hi̯-.[15] In other words, we can try a reconstruction *Hi̯eh₁k-. Unfortunately, there is no such *Hi̯eh₁k- among the known roots of Proto-Indo-

---

[13] See Létoublon (1985: 157–62) who adds the future ἵξομαι and the perfect (ἀφ)ῖγμαι to this paradigm. After a semantic comparison of the presents ἱκάνω and ἵκω she suggests that 'chaque fois qu'existe au présent une opposition morphologique entre un présent dérivé et un présent radical, le dérivé tend à prendre une valeur "effective", le radical gardant le sens duratif non-effectif'. She thus leaves the perfective meaning of the present unexplained and overlooks that ἵκω cannot be both derived from *seik- *and* a 'présent radical' (which would have to be *εἴκω).

[14] The correspondence of an active present ἵκω with a middle aorist ἱκόμην would be exceptional. For the same reason the (Homeric) middle present ἱκάνομαι must be the original counterpart of ἱκόμην and the active ἱκάνω must have been built secondarily, perhaps in analogy to the frequent (but formally unrelated) verbs in -άνω (with short -ἄ-, Risch 1974: 271–2). Note that ἱκνέομαι has no active forms although it is derived from the same original paradigm as ἱκάνω (singular *ἱκνευ-μ- => ἱκνέομαι, but plural *ἱκνϝ-ο- => ἱκάνω, Risch 1974: 271).

[15] Against Forssman (1987: 118 and 119 n. 26) and Rix (1992: 60 and 70) the development *Hi̯- > ζ- is probably not regular: cf. Lejeune (1972: 166–7), Brixhe (1979: 249–55, sociolectal?), Meier-Brügger (1992: 112), and Sihler (1995: 187–8). According to Peters (1976: 161) and J. Schindler *apud* Mayrhofer (1992–6: 2.406, s.v. *YAS*) one should even assume the inverse, i.e. Skr. *y-* = Gr. ζ- < *i̯- and Skr. *y-* = Gr. /h-/ < *Hi̯-.

European. But if we subscribe for a moment, and just for heuristic purposes, to a Benvenistian view of the basic structure of Indo-European roots, this *H_ieh_1k-* must be analysed as a 'Theme II' *H_ieh_1-* with some sort of suffixal *-k-*. And in fact there is one rather famous root *H_ieh_1-*. *H_ieh_1-* is known from Hittite *pi-yezzi* 'he sends' and *(i)ezzi* 'he does, he makes' (also Luvian *a(ya)-* 'to do, to make'), from Tocharian A *ya-* 'to do, to make', and above all of course from Greek ἵημι 'to release, to let go, to throw, to send' and from Latin *iacere* 'to throw, to cast' with the perfect *iēcī*.[16] The original meaning of the root *H_ieh_1-* may have been something like 'to release, to let go' since this accounts for 'to throw' on the one hand and for 'to make' on the other.

The root *H_ieh_1-* takes a curious and ultimately unexplained extension in *-k-* in both the Greek aorist ἧκα and the Latin perfect *iēcī* (cf. e.g. Schwyzer 1939: 741–2, Kimball 1991: 141–2). Thus, we seem to find parallels for our postulated *H_ieh_1k-* rather easily. However, there are two problems:

(1) The reconstruction of a root form *H_ieh_1k-* for Proto-Indo-European has been questioned by Untermann (1993), who points out that the *-k-* has dissimilar functions in Latin and Greek.[17]

(2) Is there any convincing way of explaining a shift from a transitive meaning 'to release' or 'to throw' in the active voice of *H_ieh_1k-* to an intransitive meaning 'to come', also in the active voice?

6. Perfects and presents: *He-H_ioh_1-ka

The second problem (2) may be tackled from *within* Greek. This has the advantage that we thereby circumnavigate difficulty (1), Untermann's reservations about the link between the two *-k-* forms in Greek and Latin: we will not have to postulate the Proto-Indo-European existence of an enlarged root *H_ieh_1k-*. Moreover, the presence of the *-k-* enlargement in the aorist ἧκα supports, but is not vital for, the solution proposed here.

As is well-known, a formation in *-k-* is also characteristic of the Greek perfect of vocalic verbal stems. A recent explanation by Kimball (1991) sees the origin of the *-k-* perfect in reanalysed aorist forms such as ἔθηκ-α (which stood

---

[16] Frisk (1960–72: 1.715, s.v. ἵημι), Chantraine (1968–80: 459, s.v. ἵημι), Melchert (1984: 159–61), Ernout-Meillet (1985: 304, s.v. *iaciō*), Walde-Hofmann (1938–56: 1.666–7, s.v. *iaciō*), Rix et al. (2001: 225, s.v. *H_ieh_1-*), Adams (1999: 492, s.v. *yām-*); critically on the Hittite connection Puhvel (1984–: 1/2.345–6, s.v. *iya-, ie-, i-*). The initial laryngeal of *H_ieh_1-* is guaranteed by the length of the initial vowel in ἵημι (Peters 1976).

[17] If the etymology of Greek νίκη 'victory' < *$ni-Hih_1k$-* 'throwing down' is correct (Thieme 1971: 382 n. 10, Klingenschmitt 1975*b*: 162 n. 22), the Greek distribution of the suffix *-k-* is wider than Untermann (1993: 461–2) admits.

beside root aorists such as reconstructed *ἔθη). This view may well be correct, but again it is ultimately irrelevant for our purposes: after all, the existence of a *-k-* suffix in the Greek perfect of stems ending in a long vowel is undeniable, no matter what its origin was. How early the *-k-* perfect came into being is difficult to say, but with ἕστηκα and βέβηκα it is certainly well-established in Homeric Greek. The use of the pluperfect βεβήκει in epic line-end formulae such as διὲκ μεγάροιο βεβήκει, οἶκόνδε βεβήκει, or Ἄιδόσδε βεβήκει even suggests a relatively early date.[18]

Despite their intransitive stative meaning, the oldest Indo-European perfect formations had only one set of endings (represented in Greek by the endings of the active perfect). An intransitive active perfect γέγονα 'I have become' belongs to an intransitive middle present γίγνομαι 'to become', and an intransitive active perfect πέποιθα 'to trust' to an intransitive middle present πείθομαι 'to get convinced', *not* to its transitive counterpart πείθω 'to convince' (cf. e.g. Szemerényi 1996: 289).

With this in mind, we may reconstruct the regular Proto-Greek perfect of the root *H̥ieh₁-. To the reduplicated root in the *o*-grade the Greek perfect marker *-k-* is added: *He-H̥ioh₁-ka.[19] The meaning of such a form must have been intransitive and corresponded to the middle present of the root *H̥ieh₁-. In Greek terms *He-H̥ioh₁-ka would have been semantically closer to the middle present ἵεμαι than to the active present ἵημι. The attested meaning of the middle (or medio-passive) ἵεμαι is 'to speed oneself, to hasten', originally *'to release oneself, to throw oneself' or *'to be released, to be thrown'. In Latin, where there is no paradigmatic middle form, the reflexive *se iacere* has exactly the same value: 'to rush'.[20]

The basic meaning of the perfect would be similar, but with an added perfective or 'stative' nuance. Thus, our reconstructed *He-H̥ioh₁-ka would have meant something like 'to have rushed somewhere' or, to make the non-past meaning of the Greek perfect clearer, 'to be somewhere as a result of rushing there'.[21] This is a rather complicated paraphrase of what we might express more

---

[18] *Il.* 6.495; 16.856; 22.362; *Od.* 1.360; 3.410; 6.11; 10.388; 17.61; 18.185; 19.47; 19.503; 20.144; 21.354; 22.433; 23.292.

[19] For the sake of clarity laryngeals are represented here and on the following pages only where their reconstruction forms part of the argument but not where their presence is irrelevant, e.g. in the perfect ending *-ka < *-kh₂e or in the present ending *-ō (if from *-oH).

[20] For a semantic parallel cf. e.g. Old Irish *fo-ruimi* 'to set, to place, to lay prostrate, to overthrow', but in reflexive and middle constructions 'to betake oneself, to go', or the frequent Old Irish perfect *dom-rala* etc. 'I came', but literally 'it threw me': *Dictionary of the Irish Language based mainly on Old and Middle Irish materials* (compact edn.), Dublin 1990, 229 and 336 (col. D235-6, s.v. *do-cuirethar* II(b), and F371-2, s.v. *fo-ruimi* (d)).

[21] Or, with an inanimate subject, 'to have been thrown/sent somewhere, to be somewhere'.

simply with 'to have come, to have reached'. The only small difference lies in the fact that 'to have come' does not say anything about the speed which was used to get into that state,[22] but this 'omission' has a good parallel outside Greek.

In Latin, the old perfective or 'stative' value of the perfect became compromised because the perfect tense incorporated both old aorist and old perfect forms. However, there was a whole class of verbs with a stative value reminiscent of the Greek perfect: the verbs in *-ēre*. For instance, the root **sed-* appears in a reduplicated present *sīdō* < **si-sd-ō* 'I sit down' with a perfect *sēdī* 'I sat down' and in a stative present formation *sedeō* 'to sit'. Hence, the *semantically* corresponding form to a Greek perfect **He-Hi̯oh₁-ka* would be a Latin verb in *-ēre* from the root **Hi̯eh₁-*. This verb is *iacēre* 'to lie'. Note that here, too, we have no indication with what speed the lying person or thing has reached that state.

Once a perfect **He-Hi̯oh₁-ka* had acquired the simple meaning 'to have come, to have reached', the semantic connection between **He-Hi̯oh₁-ka* and the ancestors of the presents ἵημι 'to release, to throw' and ἵεμαι 'to throw oneself, to rush' would have become very weak. As a consequence, the paradigmatic link between **He-Hi̯oh₁-ka* and the ancestor of ἵεμαι could be neglected and a new present for the 'orphan' perfect could be built.

At this point **He-Hi̯oh₁-ka* might have been reanalysed in analogy with regular perfects (of roots ending in a consonant) without the **-k-* suffix, not all of which have a corresponding middle present. Early epic perfects of this group include for instance ἔοργα 'to have done' beside the present **u̯rgi̯ō* (which was later remodelled into ἔργω and ῥέζω), λέλογχα 'to have obtained' beside λαγχάνω, πέπονθα 'to have suffered' beside πάσχω, τέτοκα beside τίκτω (Hesiod), or λέλοιπα 'to have left' beside λείπω (cf. Chantraine 1958: 424–5). Taking the last example as a model, the following analogical equation would operate:

$$*HeHi̯oh_1ka : x = le\text{-}loip\text{-}a : leip\text{-}\bar{o}$$

The resulting *x* would be **Hi̯eh₁k-ō*, and this would develop by regular sound changes into ἥκω. As for the meaning of the new present, the analogy would be:

---

[22] The notion of speed could be secondary anyway if the original meaning of **Hi̯eh₁-* was 'to put (down)' rather than 'to release, to throw', as may be suggested by the Tocharian and Anatolian meaning 'to do, to make' (also Greek ἥρως 'hero' < 'doing deeds, potent, capable'?, cf. Adams 1987: 176–7; for the development 'to put (down), to throw' > 'to do' Adams compares **dheh₁-* 'to put' > Latin *facere*, Engl. *do*). In this case **He-Hi̯oh₁-ka* would simply mean 'to have put oneself (into some place), to be there'.

*'I have come' : 'x' = 'I have left' : 'I leave'

Here 'x' should be 'I come'. However, if ἥκω originated as a back-formation from a perfect form, it might well retain a perfective nuance, which would explain the attested meaning 'to have come'. Interestingly Wackernagel (1925: 60) pointed out a long time ago that the perfective meaning of ἥκω must indicate an origin in the perfect (although he then simply suggested a present-type inflection of an original perfect, without arguing in more detail what kind of perfect that should have been). Typologically, we may compare the Hellenistic formation of a present στήκω 'to stand' after the perfect ἕστηκα 'to stand': again, the perfective meaning of ἕστηκα is exactly preserved.[23]

7. Problems and modifications: *He-Hi̯eh₁-ka

It must be admitted that the development proposed in section (6) is more problematic than it may seem at first sight. Most of the potential models cited there are indeed active presents with a corresponding reduplicated *o*-grade perfect, but with the exception of λείπω/λέλοιπα their present tense is not built on the simple root (but with a nasal infix, a reduplication, a *-i̯-* suffix, or a *-sk-* suffix). An analogical equation on the basis of ἔοργα, for instance, would have yielded a present *Hi̯eh₁k-i̯ō. The case of λείπω, on the other hand, is special because λείπω itself seems to be a relatively young formation replacing λιμπάνω.

Even so it may not be chronologically impossible that the analogy with the type λείπω/λέλοιπα could have worked. However, it will be shown below that we need a relatively early date in order to account for ἵκω, and middle presents such as δέρκομαι and πείθομαι with their corresponding active perfects δέδορκα and πέποιθα might rather have suggested the formation of *ἥκομαι. We might therefore retreat to the position that the creation of an active present without the addition of stem affixes simply represented the 'easiest' way of building a new present at all (no matter if there were exact analogies since the case of στήκω shows that such things do not necessarily presuppose analogical pressure). Some support for this view might be found in the occasional attempts to build a present stem εἴκω on the basis of the perfect ἔοικα, especially if the first attestation of this εἴκω is really found as early as the *Iliad* (18.520, but here

---

[23] An isolated or dialectal earlier attempt to build a present στήκω seems to be documented in Empedocles fr. 35.8 D.-K. with the imperfect (?) ἔστηκε (cf. Wright 1995: 207, Diels-Kranz 1996: 327).

εἶκε *might* belong to the paradigm of εἴκω 'to give way') and not just in classical Greek (Ar. *Nub.* 1001: fut. εἴξεις; cf. Chantraine 1968–80: 354, s.v. ἔοικα).

In spite of this possibility a slightly different explanation (along similar lines) seems preferable. It is not certain that, in positing *He-Hioh₁-ka*, we were reconstructing the correct perfect for the ancestor of ἵεμαι. Let us compare the case of the structurally similar verb ἵστημι 'to set up, to place' with the medio-passive present ἵσταμαι 'to place oneself'. The perfect corresponding to the latter is again an active formation (first attested in the Homeric poems): ἕστηκα 'to have placed oneself, to stand'. The underlying form of ἕστηκα (or Doric ἕστακα) obviously has no *o*-grade in the root but has been brought in line with the vowel colouring of the present: instead of *se-stoh₂-ka* we reconstruct *se-steh₂-ka* (cf. Kimball 1991: 147).[24]

Accordingly it may be preferable to reconstruct *He-Hieh₁-ka* as the original perfect of intransitive *Hieh₁-*. It is true that there are some isolated dialectal forms suggesting that a perfect *He-Hioh₁-ka* > (*)ἕωκα once did exist, but this *He-Hioh₁-ka* may have been a *later* formation which was and remained fully inscribed into the paradigm of ἵημι and which is therefore irrelevant in the present discussion.[25]

A perfect ἕηκα, on the other hand, is of course attested for ἵημι, but with a *transitive* meaning 'I have released'. As such it belongs to a later stage, too, and must have been recreated after the original perfect *He-Hieh₁-ka* with the old *intransitive* meaning had disappeared from the paradigm.

If we now start from an intransitive perfect *He-Hieh₁-ka* 'I have reached', we may envisage another possible way for ἥκω to have come into existence. The third person singular *He-Hieh₁-ke* 'he/she has reached' could easily be reinterpreted as an augmented imperfect form as soon as the semantic link with the ancestor of ἵεμαι had been lost and as soon as the secondary ending of the imperfect had lost its word-final consonant *-t*. Since this was presumably later

---

[24] Because of *h₂ogmós* > ὀγμός 'furrow' beside *h₂eg-* 'to drive' and similar examples, it seems unlikely that ἕστηκα can go back to *sestoh₂ka* (for a discussion of this problem see e.g. Sihler 1995: 45–6, who adopts the opposing view). The perfect τέθηκα (semantically) belongs together with τίθημι rather than τίθεμαι and must therefore be more recent; hence, it is less valuable as a parallel without *o*-grade.

[25] *He-Hioh₁-ka vel sim.* must be posited mainly because of a few sporadic *middle* perfects with a stem ἑω- in Herodotus (2.165 ἀνέωνται), in the New Testament (*Ev. Joh.* 20.23 ἀφέωνται), and in the dialects of Heracleia (*DGE* 62.153 ἀνηεῶσθαι) and Arcadia (list in Dubois 1986: 164). These 'can only have come from an active *ἕω[κα], since ω < *oh₁ would have been regular nowhere else in the original paradigm' (Kimball 1991: 146); cf. the discussion in Dubois (1986: 164-5), who also regards them as secondary formations preceded by *He-Hieh₁-ka* > *ἕηκα. The *active* form ἀφέωκα is attested even more rarely (Herodian, papyri) and may be a post-classical innovation in colloquial language (cf. Dubois 1986: 165).

than the disappearance of the laryngeals, the development would be the following:

perf. **He-Hi̯eh₁-ke* > * *ē̆-i̯ē̆-ke*   reanalysed ('impf.') **ē̆-i̯ēk-e*   new 'pres.' **i̯ēk-ō* > ἥκω

The reinterpretation of the old perfect **ē̆-i̯ē̆-ke* and the subsequent creation of a new present were probably promoted by two factors. First, there was the slightly irregular aorist stem of ἵημι with its **-k-* suffix: ἧκα < **(e-)i̯ēk-m̥*. The existence of this aorist must have supported the erroneous impression that **i̯ēk-* (instead of **i̯ē̆-*) was the verbal stem also in the perfect. Hence, in a first step the perfect **ē̆-i̯ē̆-ke* would have been misinterpreted as **ē̆-i̯ēk-e* and in a second step this would have been treated as an imperfect (with a long augment as in verbs with an initial digamma or in ἤμελλε[26]).

Second, the reanalysis was helped by the existence of early pluperfect formations with thematic secondary endings. Thus, **ē̆-i̯ē̆-k-e* could also represent the pluperfect 'he/she had reached', with a corresponding plural **ē̆-i̯ē̆-k-on* 'they had reached'.[27] Taking such cases for imperfect forms was not unnatural. Moreover, this scenario accounts for the 'perfective' meaning of ἥκω:

'impf.' **ē̆- i̯ēk-e* 'he had reached', **ē̆-i̯ēk-on* 'they had reached'
—> 'pres.' **i̯ēk-ei* > ἥκει 'he has reached' (rather than 'he reaches')

Although this explanation may seem contorted, it has a parallel elsewhere. The creation of a new (Homeric) present ἀνώγω on the basis of the (misunderstood) perfect ἄνωγε must have taken exactly this route: ἄνωγε was taken for an imperfect and ἀνώγω was built to complete the paradigm (for details see Chantraine 1958: 312, and 1968–80: 94, s.v. ἄνωγα).

8. Inventing ἵκω

When the present **i̯ēkō̄* had been built and when it had taken over the intransitive perfect semantics of **He-Hi̯eh₁-ka* > **ē̆-i̯ē̆-ka* (or, according to the variant discussed in section (6), **He-Hi̯oh₁-ka*), the latter became superfluous

---

[26] Note that the entire scenario also works if the intransitive perfect of the root **Hi̯eh₁-* was first built after the disappearance of the laryngeals. We would then start from a reduplicated perfect of **i̯ē̆-*, i.e. **i̯e-i̯ē̆-ke*. After the change of **i̯-* into /h-/, this perfect would become /(h)ehḗke/ (with 'Hauchdissimilation'), from which the new 'root' /hḗk-/ could be extracted even more easily since the apparent augment would then be a short /e-/ as in most verbs.

[27] For such (Homeric) pluperfects see Chantraine (1958: 438–9).

and was lost (though, as mentioned in section (7), both ἔωκα and ἔηκα were recreated later as *transitive* perfects of ἵημι).

In the meantime the new 'root' *Hi̯eh₁k-* or *i̯ēk-* was treated like other roots and a synonymous reduplicated present with a zero-grade vocalism was shaped: compare, for instance, Homeric μίμνω < *mi-mn-ō* 'to stay, to stand fast, to await' from the root *men-* with the synonymous unreduplicated present μένω. Starting from *Hi̯eh₁k-* the regular outcome would be *Hi-Hih₁k-* and that would ultimately lead to ἵκω with a long initial vowel. The two systems match exactly:

$$\text{μί-μν-ω} : \text{μέν-ω} \; = \; \text{*}Hi\text{-}Hih_1k\text{-}\bar{o}\; (>\text{ἵκω}) : \text{*}Hi̯eh_1k\text{-}\bar{o}\; (> \text{ἥκω})$$

If, on the other hand, the later development via *i̯ēk-* is preferable (i.e. after the laryngeals had disappeared), we must first decide what would have been the appropriate zero grade of such a recent long-vocalic root. It cannot be excluded that the long vowel would have been disregarded and that a reduplicated formation *i̯i-i̯k-* > ἱκ- would have resulted. The fact that the middle of ἵημι is ἵεμαι, not *ἶμαι, certainly does not imply that *i̯i-i̯ek-* > *ἱεκ- *must* be the expected zero grade since the long ἱ- of ἵημι/ἵεμαι suggests that the formation of ἵημι/ἵεμαι *pre*dates the loss of the laryngeals so that ἵεμαι (which *should* go back to *Hi-Hih₁-C-*) is most likely an analogical replacement of the expected *ἶμαι (the pair τίθημι/τίθεμαι perhaps acting as a model).

At the same time there is nothing to speak against a zero grade *i̯i-i̯ek-* > *ἱεκ-, which does in fact appear to be the more likely variant. Assuming that *ἱέκω would be the regular outcome of a reduplicated by-form beside ἥκω, we must take into account that an anlaut ἱε- undergoes assimilation and contraction in several dialects: the Aeolic (Lesbian) and Ionic form of (Attic) ἱερός < *i(h)erós* < *ish₁rós* is ἶρος/ἱρός. As for Doric, we might be dealing with a dialectal borrowing since in analogy with ἱερός/ἱαρός one should expect a dissimilated form *ἱάκω.[28] However, it is also possible that the semantic and phonological closeness of the unrelated *sik-* > ἱκ- prevented the opening of the second vowel and favoured the assimilation and contraction of an original *i̯i-i̯ek-*.[29]

---

[28] For such an early dialectal borrowing from Ionic into other dialect groups cf. the spread (and remodelling) of εἰρήνη 'peace' (Thess. ἰρείνα, Delph. εἰρήνα, Cretan ἰρήνα, but elsewhere in Doric ἰράνα, etc.): Meillet (1965: 230–1), Chantraine (1968–80: 324, s.v. εἰρήνη).

[29] A rather unlikely alternative would be that ἵκω was built directly as the new present of a perfect *He-Hi̯oh₁-ka*: this would presuppose that the model of τίκτω < *ti-tk-ō* beside τέτοκα was available early enough. In theory, *Hi̯eh₁-kō* > ἥκω could then be secondary after *Hi-Hih₁-kō* > ἵκω (again in analogy with cases like μίμνω/μένω).

It follows that the contracted stage of ἵκω may have been reached earlier than if ἵκω had originated from **si-sik-*. Even if we operate with **i̯i-i̯ek-ō̆*, the contraction can easily have happened before the loss of intervocalic /h/ < **-s-*, which must be post-Mycenaean. Thus we explain why there is no trace of an uncontracted ἵκω in the Homeric poems and why ἵκω does occur with one long syllable even in the frequent and archaic line-end formula (αὖτὴ δ' εἰς/σέλας δ' εἰς/κλέος) οὐρανὸν ἵκει/-ε(ν)/-η (cf. section (2) above). Note, moreover, that the subject in this formula is always inanimate, which would be puzzling if **seik-* with its original meaning 'to put forward either hand or foot, to reach out, to step' stood behind ἵκω.

Not surprisingly, in historical times when there was a short-vocalic root ἱκ- and a long-vocalic root ἱκ- with similar meanings, the two came to be associated with each other.[30] And yet, the difference of origin is still reflected in the distribution of active and middle forms pointed out in section (2): the family of **sik-* always occurs in the middle voice (except in the nasal stem ἱκάνω, which must be younger than ἱκάνομαι), that of **H̥i̯eh₁k-/*i̯ēk-* always in the active voice.[31]

In the end all the Greek dialects preferred either ἥκω or ἵκω. This was not a decision for one verbal root or another. After all, even the Attic-Ionic group which chose ἥκω continued to use **sik-* in ἀφικνέομαι and its cognates. It was rather a decision, made at a relatively late stage, for one out of two present formations. A parallel is the post-Homeric disappearance (outside poetry) of μίμνω in the pair μίμνω/μένω. Presumably the reduplicated ἵκω would have

---

[30] Also syntactically: the accusative after **H̥i̯eh₁-k-* must have been a directional accusative from the beginning, whereas the accusative after **sik-* may have originated as an object accusative (cf. Létoublon 1985: 150–1). Later both were interpreted as directional accusatives and gradually replaced by prepositional clauses.

[31] Homeric ἵκμενος, an epithet of οὖρος 'wind', should therefore be grouped together with **sik-*. The epithet is semantically obscure but presumably old (because of its athematic formation, cf. Schwyzer 1939: 751, Chantraine 1958: 384). Since winds do not have human extremities, the original meaning may be passive rather than middle (despite Létoublon 1985: 150, 'le vent favorable touche les voiles'): the wind is 'prayed for, supplicated, beseeched, (hence:) favourable' (cf. ἱκνέομαι 'to be a suppliant'). With the above explanation of ἵκω another Homeric problem finds an easy solution, too: the 'thematic sigmatic aorists' ἷξε/ἷξον. According to Prince Roth (1973: 185) ἷξε was formed when ἷκε 'came to be felt as imperfect and not aorist'. But if ἷκε has in origin nothing to do with ἱκόμην, it *is* an imperfect (of ἵκω), not the active counterpart of ἱκόμην. The creation of an aorist ἷξε was straightforward as soon as ἵκω was no longer analysable as a reduplicated present. Again this argues for an early date of the anlaut contraction (i.e. against **si-sik-*). The plural ἷξον is certainly formed after the singular ἷξε, but why should it not be ἷξαν (cf. Leumann 1953: 213)? I suspect that, when ἷξον was created, ἵκω was already taken to be a cognate of ἱκόμην so that there was the wish for the same kind of thematic aorist; but since the slot of ἷκον was taken up by the imperfect, one had to content oneself with the closest possible approximation.

been doomed similarly in all dialects if it had not had the powerful support from the family of ἰκνέομαι. In any case, in order both to account for ἵκω and ἥκω and to maintain a formally strict view of phonological reconstruction we no longer need to postulate a parallel survival since Proto-Indo-European times of two semantically identical but etymologically unrelated verbs.

9. Conclusion

The more general point of this paper was to show that our attitude towards semantics in reconstruction and etymology considerably influences the results we get. In my explanation of ἵκω and ἥκω I have opted for conservative and cautious views on Indo-European historical phonology — so to speak, for the greatest common denominator of what the standard handbooks teach. At the same time I firmly believe that semantics is no *quantité négligeable*. The price to pay for such a conciliatory attitude in the conflict of semantic and formal interests is often — and in this paper no less than elsewhere — the need to assume and try to defend complicated historical developments. If the resulting theories fail to convince everyone, they may at least remind us of the fact that it is sometimes little more than a question of taste 'how much semantics we take with our phonology'.

References

Adams, Douglas Q. (1987). '῝Ηρως and ῝Ηρα: of men and heroes in Greek and Indo-European'. *Glotta*, 65: 171–8.
— (1999). *A Dictionary of Tocharian B*. Amsterdam/Atlanta.
Benveniste, Émile (1962). *Origines de la formation des noms en indo-européen*. Paris.
— (1969). *Le vocabulaire des institutions indo-européennes: 2. Pouvoir, droit, religion*. Paris.
Brixhe, Claude (1979). 'Sociolinguistique et langues anciennes: A propos de quelques traitements phonétiques irréguliers en grec'. *BSL*, 74: 237–59.
Chantraine, Pierre (1958). *Grammaire homérique I: Phonétique et morphologie*. Paris.
— (1968–80). *Dictionnaire étymologique de la langue grecque: Histoire des mots*. Paris.
Diels, Hermann, and Kranz, Walther (1996). *Die Fragmente der Vorsokratiker I* (19th edn. = 6th edn., 1951). Zürich.

Dubois, Laurent (1986). *Recherches sur le dialecte arcadien I: Grammaire* (BCILL, 33). Louvain-la-Neuve.

Ernout, Alfred, and Meillet, Antoine (1985). *Dictionnaire étymologique de la langue latine: Histoire des mots* (4th edn. by Jacques André). Paris.

Forssman, Bernhard (1987). 'Vedisch *āyavasa-*', in George Cardona and Norman H. Zide (eds.), *Festschrift for Henry Hoenigswald on the Occasion of his Seventieth Birthday*. Tübingen, 115–19.

Fraenkel, Ernst (1962–5). *Litauisches etymologisches Wörterbuch*. Heidelberg/ Göttingen.

Frisk, Hjalmar (1960–72). *Griechisches etymologisches Wörterbuch*. Heidelberg.

García Ramón, José Luis (1999). 'Zur Bedeutung indogermanischer Verbalwurzeln: *$h_2$nek-* "erreichen, reichen bis", *$h_1$nek-* "erhalten, (weg)nehmen"', in Jürgen Habisreitiger, Robert Plath, and Sabine Ziegler (eds.), *Gering und doch von Herzen: 25 indogermanistische Beiträge Bernhard Forssman zum 65. Geburtstag*. Wiesbaden, 47–80.

Hamp, Eric P. (1982). 'Indo-European notes'. *IF*, 87: 70–81.

Kimball, Sara E. (1991). 'The origin of the Greek κ-perfect'. *Glotta*, 69: 141–53.

Klingenschmitt, Gert (1975*a*). 'Altindisch *śáśvat-*'. *MSS*, 33: 67–78.

— (1975*b*). 'Tocharisch und Urindogermanisch', in Helmut Rix (ed.), *Flexion und Wortbildung: Akten der V. Fachtagung der Indogermanischen Gesellschaft*. Wiesbaden, 148–63.

— (1982). *Das altarmenische Verbum*. Wiesbaden.

Lamberterie, Charles de (1990). *Les adjectifs grecs en -υς: Sémantique et comparaison I*. Louvain-la-Neuve.

Lejeune, Michel (1972). *Phonétique historique du mycénien et du grec ancien*. Paris.

Létoublon, Françoise (1985). *Il allait, pareil à la nuit: Les verbes de mouvement en grec — supplétisme et aspect verbal*. Paris.

Leumann, Manu (1953). '"Aoristi mixti" und Imperative vom Futurstamm im Griechischen'. *Glotta*, 32: 204–53.

Martinet, André (1964). *Économie des changements phonétiques: Traité de phonologie diachronique*. Bern.

Mayrhofer, Manfred (1986). *Indogermanische Grammatik I.2: Lautlehre [Segmentale Phonologie des Indogermanischen]*. Heidelberg.

— (1992–6). *Etymologisches Wörterbuch des Altindoarischen*. Heidelberg.

Meier-Brügger, Michael (1987). 'Griech. ἐνεῖκαι und ἐνεγκεῖν: vereinigen oder trennen?'. *ZVS*, 100: 313–22.

— (1992). *Griechische Sprachwissenschaft II: Wortschatz, Formenlehre, Lautlehre, Indizes*. Berlin/New York.

Meillet, Antoine (1965). *Aperçu d'une histoire de la langue grecque* (8th edn. with revised bibliography by Olivier Masson). Paris.

Meiser, Gerhard (1986). *Lautgeschichte der umbrischen Sprache*. Innsbruck.

Melchert, H. Craig (1984). *Studies in Hittite Historical Phonology*. Göttingen.

Morpurgo Davies, Anna (1998). *History of Linguistics IV: Nineteenth-Century Linguistics*. London/New York.

Panagl, Oswald (1976). 'Zur Etymologie von griech. ἴχνος', in Anna Morpurgo Davies and Wolfgang Meid (eds.), *Studies in Greek, Italic, and Indo-European Linguistics Offered to Leonard R. Palmer*. Innsbruck, 273–81.

Peters, Martin (1976). 'Attisch *hī̆ēmi*'. *Sprache*, 22: 157–61.

Prince Roth, Catherine (1973). 'Thematic s-aorists in Homer'. *HSPh*, 77: 181–6.

Puhvel, Jaan (1984–). *Hittite Etymological Dictionary*. Berlin/New York/Amsterdam.

Risch, Ernst (1974). *Wortbildung der homerischen Sprache* (2nd edn.). Berlin/New York.

Rix, Helmut (1992). *Historische Grammatik des Griechischen: Laut- und Formenlehre* (2nd edn.). Darmstadt.

—, Kümmel, Martin, Zehnder, Thomas, Lipp, Reiner, and Schirmer, Brigitte (2001). *Lexikon der indogermanischen Verben: Die Wurzeln und ihre Primärstammbildungen* (2nd edn.). Wiesbaden.

Ruijgh, Cornelius J. (1957). *L'élément achéen dans la langue épique*. Assen.

— (1968). 'A propos d'une nouvelle application de méthodes structuralistes à la langue homérique'. *Mnemosyne* 4th ser., 21: 113–31.

Schmid, Wolfgang P. (1956). 'Vedisch *tmahe* und Verwandtes'. *IF*, 62: 219–39.

Schwyzer, Eduard (1939). *Griechische Grammatik I: Allgemeiner Teil, Lautlehre, Wortbildung, Flexion*. München.

Sihler, Andrew L. (1995). *New Comparative Grammar of Greek and Latin*. New York/Oxford.

Strunk, Klaus (1988). 'Über Laryngale und einige reduplizierte Verbalstämme', in Alfred Bammesberger (ed.), *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*. Heidelberg, 563–82.

Szemerényi, Oswald J. L. (1996). *Introduction to Indo-European Linguistics*. Oxford.

Thieme, Paul (1971). 'Beseelung in Sprache, Dichtung und Religion', in Paul Thieme, *Kleine Schriften I*. Wiesbaden, 374–85.

Untermann, Jürgen (1993). 'Gr. ἔθηκα = lat. *feci*, gr. ἧκα = lat. *ieci*?', in Gerhard Meiser (ed.), *Indogermanica et Italica: Festschrift für Helmut Rix zum 65. Geburtstag*. Innsbruck, 461–8.

Wackernagel, Jacob (1925). 'Griechische Miszellen'. *Glotta*, 14: 36–67.

Walde, Alois, and Hofmann, Johann Baptist (1938–56). *Lateinisches etymologisches Wörterbuch* (3rd edn.). Heidelberg.

Willi, Andreas (2001). 'Lateinisch *iubēre*, griechisch εὐθύς und ein indogermanisches Rechtskonzept'. *HS*, 114: 117–46.

Wright, M. Rosemary (1995). *Empedocles: The Extant Fragments* (2nd edn.). London.

The 'iterative-intensives' in -σκον

*Jason Zerdin**

qitsuk qinngasaarnoqaraangami miaartortarpoq
'if you tease a cat, it meows'
West Greenlandic Eskimo (Dahl 1995: 422)

Introduction

This paper concerns the Ancient Greek forms which show the suffix -σκον. Historically, this suffix is generally held to be ultimately related to Greek present-forming -σκω and so to Indo-European *-sk̂e/o-. Synchronically, -σκον presents a semantic problem: forms which show it are generally cited as denoting repeated action, and on this account are called 'iteratives'. In some instances, however, they do not indicate repeated action, and on these occasions scholars use terms such as 'intensive' and 'durative' to describe them. This has led to the forms in -σκον often being referred to with the joint term 'iterative-intensive': a description of the problem, but not a solution. An inquiry into the distribution of the two meanings can in fact show that the answer lies in the meaning of the base forms from which the forms in -σκον are derived.

*1. Indo-European *-sk̂e/o- and Greek -σκω and -σκον*

*1.1. Indo-European *-sk̂e/o-*

As mentioned above, the Indo-European ancestor of both -σκω and -σκον is generally reconstructed as present-forming *-sk̂e/o-, which is found in all major branches of the Indo-European language family.[1] Two examples are:

---

[1] On this suffix see especially Brugmann and Delbrück (1893-1916: II/3.350-61, IV.59-64), Porzig (1927: 154-67), Watkins (1969: 56-8, 111), Keller (1992: 167-97 on *-sk̂e/o-, 279-92 on *-isk̂e/o-); Szemerényi (1996: 273-4), and Rix (1998: 19). Most scholars no longer reconstruct a separate Indo-European preform *-isk̂e/o-: cf. Klingenschmitt (1982: 73-7), Ruijgh (1985: 140, 148-9), Vine (1993: 49-60), but contrast Keller (1992: 279-81, 287-92).

(1) (a) Proto-Indo-European $g^w m̥-sḱe/o-$ > Vedic *gácchati* 'go, come'; Greek βάσκω 'go'; cf. also Young Avestan *jasaiti* 'go, come' and perhaps Tocharian A *kumnäṣ*, B *känmaṣṣäm* 'come' (cf. Rix 1998: 187-8);

(b) Proto-Indo-European *$*prḱ-sḱe/o-$ > Vedic *pṛccháti* 'ask', Old Avestan *pərəsa-* 'ask', Old Persian *pṛsāmiy* 'punish'; Armenian aorist *eharcʻ* 'ask'; Latin *poscō* 'demand, ask'; Old Irish *-airc*, Middle Welsh *eirch* 'beseech, ask'; cf. also Oscan future perfect *comparascuster* 'consulta erit' and perhaps Old High German *forscōn* 'search' (cf. Rix 1998: 442-3).

1.2. Greek -σκω and -σκον

Forms in Greek -σκω[2] and -σκον[3] differ in several ways. Although both types are found in Homer, only the former becomes at all productive later on; the latter is rare in post-Homeric Greek, only occurring with any degree of frequency in Herodotus. There are notable morphological and semantic differences between the two sets.

1.2.1. Morphology

Greek -σκω is one of many Greek present-forming suffixes. The basic types of formation are:[4]

---

[2] There is not a great deal of bibliography on present-forming Greek -σκω. See in general Schwyzer and Debrunner (1939-71: 1.707-12), Chantraine (1991: 223-7), Rix (1992: 213-14), and Sihler (1995: 505-7). In works on the Greek verb, cf. Ruipérez (1954: 130-5), Kujọṛẹ (1973: 7-8, 92-7, on Homer, 251-9 on Classical Greek, 296-7 on post-Classical Greek), and Duhoux (1992: 335-7). Works on -σκω alone (both also touching on -σκον) are Giacalone Ramat (1967) and Zerdin (2000). There are also useful summaries in Ruijgh (1985: 141-5) and Keller (1992: 179-84 on -σκ-, 282-3 on -ισκ-).

[3] There is much more bibliography on the subject of the forms in -σκον than there is on presents in -σκω. On the formation in general see Curtius (1852: 27-9, 1877-80: 2.405-13), Brugmann and Delbrück (1893-1916: II/3.357, IV.62-4), Bechtel (1921-4: 1.85, 3.215), Schwyzer and Debrunner (1939-71: 1.710-12), Meillet and Vendryes (1948: 242-3), Chantraine (1986-8: 1.318-25; and on this Shipp 1972: 88-90), Chantraine (1991: 226-7), Ruipérez (1954: 132-4), Rosén (1962: 125-6), Bottin (1969: 116-24), Giacalone Ramat (1967: 115-22), Risch (1974: 276-8), Ruijgh (1985: 145-8), Rix (1986: 19, 1992: 229), Duhoux (1992: 336), Meier-Brügger (1992: 1.128), Keller (1992: 181-2), Rijksbaron (1994: 15), and Sihler (1995: 506). Specific studies on this subject are Stolpe (1849), Týn (1859), Brugmann (1902), Kluge (1911: 22-61), Fantini (1950), Wathelet (1973), Negri (1976), and Kimball (1980). The formation is compared with possible relatives in other languages by Puhvel (1991: 13-20), on Hittite, and Pisani (1959: 176-7) and Clackson (1994: 75-83), both on Armenian.

[4] These data and categories are taken from Zerdin (2000); cf. e.g. the summaries of types of forms found in Homer (134-6) and Plato (359-60).

(2) (a) Root formations, of which there are no new creations after the Homeric period: βάσκω 'go' (*gʷm̥-sḱe/o-*), θνη(ί)σκω 'die' (**dhn̥h₂-sḱe/o-*);

   (b) Reduplicated formations, only formed from monosyllabic stems: Homeric γι-γνώ-σκω 'know' (apparently ultimately from **ǵn̥h₃-sḱe/o-*), τι-τρώ-σκω 'wound';

   (c) Forms in -ίσκω, generally found beside non-present stems in -η- or -ω-: ἁλ-ίσκομαι (perfect ἑάλ-ω-κα) 'be taken', εὑρ-ίσκω (perfect ηὕρ-η-κα) 'find';

   (d) Polysyllabic stems, which show the same stem before -σκω as in other tenses, and are almost entirely post-Homeric: γηρά-σκω (Classical aorist ἐγήρᾱσα) 'become old', μεθύ-σκω (aorist ἐμέθυ-σα) 'make drunk'.

The formation of past tense forms in -σκον is very different, as we shall see in detail below (§3). To summarize, they are imperfect indicatives created from present and aorist stems: e.g. σπένδ-ε-σκον from present σπένδ-ω 'make libation' beside σπείσα-σκε from its aorist ἔ-σπεισα.

1.2.2. Semantics

The forms in -σκον also form a separate group semantically. The basic difference is that -σκον usually denotes repeated action, whereas -σκω never does. For example, one instance of an imperfect of a present in -σκω from Homer is:[5]

(3) ἀλλ' ὅτε δὴ <u>γίγνωσκε</u> θεοῦ γόνον ἠὺν ἐόντα, ... (*Il.* 6.191)

But when he [the King of Lycia] <u>perceived</u> that he [Bellerophon] was the noble child of a god...

The function of -σκω in fact at first sight appears to be threefold: causative, inchoative, or zero. These three meanings also occur for descendants of **-sḱe/o-* in three of the four languages where it is productive; hence Table I (4), where glosses are taken/adapted from LSJ.[6]

---

[5] The text used in this article is that of the Oxford Classical Text of Homer's works: for the *Iliad*, Monro and Allen (1920), and for the *Odyssey*, Allen (1917-19). Glosses are taken (and sometimes adapted for the sake of brevity) from Cunliffe (1924).

[6] I argue elsewhere (Zerdin 1998, in an updated form Zerdin forthcoming) that Greek present-forming -σκω was used with inchoative meaning beside an unmarked present — either as an active, or as a medio-passive with causative active beside it. Hence e.g. inchoative γενειάσκω 'begin to get a beard' : stative γενειάω 'have a beard' and causative μεθύσκω 'make drunk' : inchoative μεθύσκομαι 'become drunk': stative μεθύω 'be drunk'. Otherwise the suffix has zero meaning, as in e.g. ἀποδιδρᾱ́σκω 'run away'.

(4)   Table I

	With -σκ-	Without -σκ-	Other I.-E. language
Causative	μεθύσκω 'make drunk, intoxicate'	μεθύω 'be drunken with wine'	Tocharian[7]
Inchoative	γενειάσκω 'begin to get a beard'	γενειάω 'get a beard; have a beard'	Latin[8]
Zero	ἀρέσκω 'please'	aorist ἤρεσα 'pleased'	Armenian[9]

Interestingly, the meaning of -σκον is found for the one further productive descendant of *-sk̂e/o-: Hittite iterative -šk-[10] — perhaps evidence that Indo-European *-sk̂e/o- also had iterative meaning.[11]

One example, formed from ἀποστρέφω 'turn back', is:

(5)   ὁσσάκι δ' ὁρμήσειε πυλάων Δαρδανιάων
       ἀντίον ἀΐξασθαι ἐϋδμήτους ὑπὸ πύργους, ...
       τοσσάκι μιν προπάροιθεν <u>ἀποστρέψασκε</u> παραφθὰς
       πρὸς πεδίον (*Il*. 22.194-5, 197-8)
       As often as he [Hector] moved to rush straight for the Dardanian gates to the well-
       built walls... so often he [Achilles], having got ahead, <u>would turn</u> him <u>back</u> towards
       the plain.

The suffix -σκον may be added to a form already in -σκω, as in:[12]

---

[7] See Adams (1988: 75-7), Pinault (1989: 125-6, 139-41), Keller (1992: 189-90), and Hackstein (1995, especially 2-4, 167-353).

[8] See Ernout (1953: 132-3), Leumann, Hofmann, and Szantyr (1965-79: 1.535-9, 2.298), Allen (1980), Keller (1992), and Haverling (2000).

[9] See Klingenschmitt (1982: 60-84) and Keller (1992: 185-6, 284).

[10] See Bechtel (1936), Kronasser and Neu (1966-87: 1.575-88), Dressler (1968: 157-236), Oettinger (1979: 315-29), and Keller (1992: 186-9, 285-7).

[11] Hittite of course gives us the oldest evidence for inherited *-sk̂e/o-; further, the Greek forms in -σκον may well be an archaism — for example, they are mostly restricted to Homer. By contrast, the Tocharian causative and Latin inchoative meanings are later developments: for example, the most ancient Latin forms, such as Latin *poscō* 'ask' (from *pr̥k̂-sk̂e/o-; cf. (1b) above), show no 'inchoative' meaning. The whole matter has been the subject of much debate and unfortunately is outside the scope of this paper; cf. the summary at Zerdin (2000: 50-9, followed up at 323-6 and 458, 466-9). Also relevant is the subject of the origin of the forms in -σκον: cf. Zerdin (2000: 281-6).

[12] However, this only seems to occur for verbs where the present formant -σκ- has been reanalysed as part of the stem: βόσκω 'feed' has future βοσκ-ή-σω (*Od*. 17.559) and μίσγω 'mix together', which shows -σκον forms μισγέσκετο (*Od*. 18.325) and ἐμισγέσκοντο (*Od*. 20.7), has e.g. the nominal μισγ-άγκεια 'place where two glens join their streams' (*Il*. 4.453).

(6)   οὐ γὰρ τῆλε νεὸς κυανοπρῴροιο
      <u>βοσκέσκονθ'</u> ἕλικες καλαὶ βόες εὐρυμέτωποι (*Od.* 12.354-5)
      for the fair cattle, with curved horns and broad brows, <u>used to graze</u> not far from our
      dark-prowed ship.

The sense of the verbs is generally held to be 'iterative' or 'intensive'.[13] As opposed to the previous iterative examples (5) and (6), an instance of the latter type is ναιετάασκε 'used to live', where no repetition is indicated — as when Odysseus says to Eumaeus:

(7)   ἠὲ διεπράθετο πτόλις ἀνδρῶν εὐρυάγυια,
      ἦ ἔνι ναιετάασκε πατὴρ καὶ πότνια μήτηρ...; (*Od.* 15.384-5)
      Was a city of men, with wide streets, sacked, in which your father and lady mother
      <u>used to live</u>...?

Hence the forms in -σκον are usually called 'iterative', 'intensive' ('durative') or the combined 'iterative-intensive'. The two main views to explain these data are as follows. Firstly, Ruipérez (1954: 130-5) argues that -σκ- is clearly durative: the frequent iterative meaning is a realization of the durative one. In this way he puts his emphasis on the so-called 'intensives', such as the aforementioned ναιετάασκε 'used to live'. On the other hand, Giacalone Ramat (1967: 106-7) claims that the iterative sense is primary; intensive sense occurs for -σκον with verbs expressing feelings and desires, e.g. φιλέεσκον 'used to love' and ποθέεσκε 'used to desire'. However, in ναιετάασκε we have neither a feeling nor a desire; so why is this form an intensive? There should be some reason for the distribution, as the concept that the suffix might randomly show different meanings with different verbs is unsatisfactory.

Do we therefore explain the iteratives as 'secondary' on the intensives, or the intensives as a subset of the iteratives? Is -σκον a marker of durativity or of iterativity? And why does it show two meanings?

2. General features

Iteratives in -σκον show secondary, indicative endings, i.e. are conjugated as if they are imperfects. The forms found in Homer almost never show the

---

[13] For 'intensive' meaning cf. Kluge (1911: 28-9), Giacalone Ramat (1967: 115ff.), Negri (1976: 236), and Clackson (1994: 76 and 215 n. 80).

augment.[14] Homer uses any person, singular or plural (e.g. ἔχεσκες, νικάσκομεν), although the third person singular is very much the most frequent.[15] As has already been noted (§1.2.2 above), the forms are either iterative or intensive in meaning.

The iteratives are nearly always formed off base verbs that are found in Homer.[16] Many of these base verbs are poetic and indeed peculiar to epic.[17] The starting point is usually a present stem (as κτείνεσκον), though it may be (more rarely) from an aorist one (ἴδεσκον): Homer shows 126 different forms, of which 31 are on aorist stems; the total number of occurrences is 250, of which 140 are found in the *Iliad* and 110 in the *Odyssey*.[18] Wathelet (1973: 393-4) lists 69 different iteratives which occur in Homer as only being found once each. As already seen, the suffix -σκον is occasionally found added to verbs which already have a -σκ- affix. Most forms in -σκον are simples, although compounds do occur.[19] In five cases, one verb gives forms from both its present and aorist stems.[20] Three pairs of cognate forms also occur.[21]

---

[14] The one major exception is (ἔ)φασκον, on which see note 18 below. Generally, wherever the augment is found, the verb may be read as augmentless without spoiling the metre, or there are textual variants, or both; cf. Wathelet (1973: 385-6 and 386 n. 43) and also, on the subject of the exception ἐμισγέσκοντο (*Od.* 20.7), Clackson (1994: 76 and 215 n. 81).

[15] Cf. Bottin's figures (1969: 117-18) — although they are not correct, as comparison with Kluge's lists at (1911: 22-6, 30-2), and also with Stolpe (1849: 7), shows.

[16] Cf. Risch (1974: 277).

[17] Cf. further Wathelet (1973: 389-90), who cites various forms and analyses their dialectal origins. For parallels between expressions where the same verbal stems appear with and without -σκον cf. Wathelet (1973: 397-9).

[18] Data were adapted from the list given in Risch (1974: 277-8), whose only omission seems to be ἴδεσκον; citations were taken from Gehring (1891). These were checked against the data given by Kluge (1911: 22-6, 30-2), Bottin (1969: 121-4), and Chantraine (1986-8: 1.321-5). See further the morphological analysis in §3 below. For criticisms of Bottin's (1969: 121-4) data see Clackson (1994: 215 n. 78). In these figures, forms in καλε(ε)- are counted as one verb (on which see further note 30 below) and three controversial forms are not included at all: παρέβασκε (*Il.* 11.104), ἔσκον (22 times in the *Iliad*, 23 in the *Odyssey*), and (ἔ)φασκον (twice in the *Iliad*, fifteen times in the *Odyssey*). The reasons for the difficulties are in all three cases different. For παρέβασκε, we may either have an iterative in -σκον, meaning 'used to go beside', or alternatively an imperfect form of βάσκω 'was going beside'. ἔσκον may be the form in -σκον from εἰμί 'be', or it may be ultimately from an inherited *h_1s-ske/o-, having possible cognates in Old Latin, Tocharian, and Palaic. Lastly, (ἔ)φασκον from φημί 'say' is the only form in -σκον which is used in later Greek: parts in φασκ- come to be used as suppletives for some of those of φημί. See further Zerdin (2000: 142-3, on παρέβασκε, and 309-23, on ἔσκον and (ἔ)φασκον).

[19] As e.g. ἐπιρρήσεσκον 'used to thrust home' (*Il.* 24.454, 456) beside ῥήσσοντες 'beating time' (*Il.* 18.571); cf. Wathelet (1973: 389).

[20] These, as noted by Delbrück (Brugmann and Delbrück 1893-1916: IV.63) and Risch (1974: 276) are ἱστάσκε/στάσκε, σπένδεσκον/σπείσασκε, φαινέσκετο/φάνεσκε, φεύγεσκε/

The forms in -σκον are also well-known to show a strong tendency to occur in series of, usually, two or three forms; differing lists of roughly twenty examples for the *Iliad* and ten for the *Odyssey* are given by Bottin (1969: 118-21), Wathelet (1973: 394-5), and Puhvel (1991: 14-16), although the phenomenon has also been said to be more frequent in later books.[22] This grouping is non-formulaic in nature.[23] Clackson (1994: 78) claims that almost half of all occurrences of iteratives are found within three lines of another iterative; in fact, all such groups together only account for 36 instances of the 138 in the *Iliad*, and for 20 of the 109 in the *Odyssey*. He points out that all but six -σα- forms are found in conjunction with other iteratives.[24]

Homeric iteratives in -σκον are therefore generally augmentless third person indicatives, with a tendency to occur in series. According to Bottin (1969, especially 116-24), these facts are all connected: the meaning of the iteratives lends them naturally to narrative, and it is in narrative passages that Homer usually omits the augment most, whilst he tends to use it with verbs which describe general actions. Hence their use in the past indicative, describing 'real' past actions. The tendency for neologisms may be due to the extended use of the suffix in series.[25]

Attestations in authors other than Homer are rare and usually in imitation of epic;[26] the one exception is Herodotus, whose usage (both morphologically and, as we shall see, semantically) is more limited than Homer's. This implies that the formation is peculiar to Ionic, as is generally accepted — unless Herodotus is epicizing when he uses the forms.[27] In Herodotus, nearly all forms are created from present stems; iteratives from aorists are only derived from thematic stems, indicating that Homer's use of iteratives from aorists in -σα (and, on one

---

φύγεσκε, and ὤθεσκε/ὦσασκε. On the difference between the aorist and present types, which apparently mirrors that of non-iterative forms, cf. Týn's conclusion (1859: 694-5), Brugmann and Delbrück (1893-1916: IV.62), Kluge (1911: 26), Chantraine (1986-8: 1.323-4), Ruijgh (1985: 146 n. 147), and Rijksbaron (1994: 15).

[21] The pairs are ναίεσκον/ναιετάασκον, ὀλέκεσκε/ἀπολέσκετο, and οὔτασκε/οὐτήσασκε.

[22] Cf. Chantraine (1986-8: 1.325), followed by Clackson (1994: 78 and 215 n. 82). The former notes especially the instances at *Il.* 24.11-24 and *Od.* 11.585-600.

[23] Cf. Wathelet (1973: 393-5).

[24] Cf. Clackson (1994: 77).

[25] Cf. Bottin (1969: 118).

[26] Cf. the data discussed by Curtius (1877-80: 2.406) and Wathelet (1973: 391-2); the morphological categories given by Curtius (1877-80: 2.409-12), Schwyzer (Schwyzer and Debrunner 1939-71: 1.711), and Fantini (1950: 320-30) contain further forms from other authors, and some dialectal forms are given by Bechtel (1921-4: 1.85, 3.215). Usage in Hesiod and the Homeric Hymns is discussed by Kluge (1911: 40-1).

[27] Thus Brugmann (1902: 270-1), disputed by Schwyzer (Schwyzer and Debrunner 1939-71: 1.710 n. 9); cf. also Kluge (1911: 43).

occasion, in -ην) was probably a peculiarly epic innovation.[28] Only the third person is used, and base forms are never athematic.[29] Here also, the forms are used to narrate past actions, and are often used in groups.

3. Morphological analysis

Before semantic analysis, a brief morphological review should prove to be of use. The following breakdown is based on that of Chantraine (1986-8: 1.321-5) and Clackson (1994: 76-8); comparable also is that of Schwyzer (1939-71: 1.711).

3.1. Present stem formations

We find 95 different iteratives formed from present stems in Homer, giving a total of 173 instances; they fall into four categories.

Firstly, there is a small group of forms derived from athematic verbs with zero-grade of the stem vowel; hence e.g. ἵστασκε from ἵστημι. There are six such formations in Homer, not including καλέσκετο;[30] all bar κέσκετο occur only once, giving seven instances in total.[31]

Iteratives on thematic stems are derived by keeping the thematic vowel -ε- (never -ο-) of the present stem and adding the suffix -σκ- and secondary endings; hence μέν-ε-σκον from μέν-ω. This is the largest category: 59 different forms occur, giving 101 instances in total.[32]

Thirdly, there are contracted presents. Iteratives of -έω and -άω verbs are quite frequently found, with one case in -όω: σώεσκον. In some cases, the final vowel of the stem is doubled (ἰσχανάασκον, ποθέεσκε) and in others it is not (νικάσκομεν, μυθέσκετο). Wathelet (1973: 387-9) ascribed the difference to a combination of factors (occurrence of contraction of vowels and diectasis, and existence of athematic forms). As is often observed,[33] the ultimate reason seems however to be metrical: the disyllabic version of the suffix follows a light syllable, and the monosyllabic version a heavy one: hence φιλέεσκον beside

---

[28] Cf. Wathelet (1973: 385 n. 42).

[29] On iteratives in Herodotus, see Stolpe (1849: 44-50), Kluge (1911: 42-8), Rosén (1962: 125-6), Bottin (1969: 116-17), whose list of forms is, however, unreliable, and Wathelet (1973: 392-3). See further §5.3 below.

[30] Oddly, καλέσκετο occurs beside a form καλέεσκον; see further Clackson (1994: 79) on the problem presented by the *seṭ* root here.

[31] For the other forms see Chantraine (1986-8: 1.322) and Clackson (1994: 76).

[32] Cf. the list at Chantraine (1986-8: 1.322).

[33] Cf. Stolpe (1849: 4), Shipp (1972: 89), Chantraine (1986-8: 1.322-3), and Kimball (1980), followed by Clackson (1994: 76).

ὤθεσκε.[34] The series -αασκ- is presumably from *-αεσκ- by diectasis.[35] We find 26 different forms from contracted verbs in Homer (counting ἔασκε together with εἴασκον, and καλέσκετο with καλέεσκον), giving a total of 59 instances of such forms. Six and five occur in -αα- and -α- respectively (totalling eleven occurrences for each type), nine and seven in -εε- and -ε- (26 and 10 occurrences, including καλέσκετο), and there is also σώεσκον.[36]

Lastly, the suffix -ασκον, as found in both iteratives from -σα- aorists (§3.2 below) and from contracted stems, is itself extended to three other stems which are generally now classed with present-stem formations: ἰσάσκετο from ἰσάζω, κρύπτασκε from κρύπτω and ῥίπτασκον from ῥίπτω, ῥιπτάζω.[37] The latter iterative is found five times, the other two only once each.

3.2. Aorist stem formations

Iteratives from aorist stems form a much smaller category: there are only 31 different forms, totalling 77 different instances; of these, nearly half are accounted for by εἴπεσκε. Two categories of these forms are parallel to the first two types of present-stem formations listed in §3.1 above. Four athematic formations give a total of nine occurrences, as δύ-σκον from ἔ-δῡ-ν.[38] Secondly, there are also seven thematic forms, as in φύγ-ε-σκον from ἔ-φυγ-ον; with εἴπεσκον occurring 28 times, these total 38 occurrences in all.[39]

In addition, Homer[40] also has iterative formations for the other types of Greek aorist. Iteratives of sigmatic aorists may be derived by adding -σκον to the aorist stem, inclusive of the -α- which occurs in all forms but the third person singular: thus θρέξα-σκον from ἔθρεξα. Nineteen such forms occur, giving 26 instances, of which all but six are found in the *Odyssey*.[41] Lastly, there is one instance of a

---

[34] The exception is καλέσκετο beside καλέεσκον; see note 30 above.

[35] Cf. Kimball (1980: 46) and Clackson (1994: 76-7).

[36] For lists of forms in -ε-, -αα-, and -α- see Chantraine (1986-8: 1.323). The figures given here do not agree with those of Clackson (1994: 76-7), who states that there are fourteen stems and 30 forms in -(έ)εσκον, and eleven stems and 24 forms in -(ά)ασκον.

[37] Cf. Curtius (1877-80: 2.410-11), Chantraine (1986-8: 1.323, 324), and Clackson (1994: 77). Risch (1974: 277) is less certain whether these forms are to be categorized as showing present or aorist stems. For further discussion of their origins see Wathelet (1973: 390).

[38] Including οὔτασκον (*Il.* 15.745), on which see Chantraine (1986-8: 1.325), who also lists the other athematic aorist-stem forms.

[39] For a list see Chantraine (1986-8: 1.324).

[40] Outside Homer, the only such example is apparently ἀλδήσασκε at Orphica *L.* 370.

[41] Cf. Chantraine's list (1986-8: 1.324-5); he omits ὤσασκον, however.

form in -σκον for the aorist passive: the presumably analogical φάν-ε-σκον,[42] taken from ἐφάνην, found four times.

4. Aristotelian aspect and generic sentences

The general impression given by the literature discussed so far is that the iteratives in -σκον are in fact a category with little semantic coherence. Before we tackle this problem, it may be worthwhile to turn to some more general considerations; in particular, we must discuss 'Aristotelian aspect'. Also known as Vendlerian aspect, this is only part of the larger category of aspect in general. The latter is a notoriously complex concept: Comrie states that 'aspects are different ways of viewing the internal temporal constituency of the situation' (1976: 3; cf. also Bybee 1992: 145), but matters are much more complicated than this, and a full discussion of these and other views is given by Binnick (1991: 207-14), who concludes that the category should in fact be divided into three parts — of which Aristotelian aspect is one.[43]

Aristotelian aspect is a lexical phenomenon, which involves the categorization of situations into categories as described by Vendler.[44] Situations[45] are first of all divided into states (non-dynamic; e.g. 'desire, want, love') and non-states (dynamic); the latter are divided into activities (atelic;[46] e.g. 'run, swim') and performances (telic); and the latter are further subdivided into achievements (momentaneous; e.g. 'recognize, find, win the race, die') and accomplishments (durative; e.g. 'run a mile, walk to school, grow up'). These categories may be neatly represented in a tree diagram, with examples given below each type (Table II).[47]

---

[42] On this form see Chantraine (1986-8: 1.325, with addendum at 1.517) and Wathelet (1973: 390).

[43] Cf. in general the main body of Binnick's monograph on tense and aspect (1991: 135-214), and his summary of his own conclusions (1991: 456-8). In this model, Aristotelian aspect is differentiated from aspect itself, which denotes the imperfective/perfective distinction; and the Aktionsarten, which here concern the starting, stopping, and pausing of the action.

[44] Cf. Binnick (1991: 143-4, 170-97, also 207-14, summarized at 457).

[45] We are here dealing not with isolated verbs, but with verb phrases; if we do not draw this distinction, then we cannot differentiate between the Aristotelian aspect of, for example, 'run' and 'run a mile'.

[46] See Binnick (1991: 189-97) for a discussion of telicity (which he calls 'telicness'); according to his definition, it does not involve 'the achievement of a goal, not the potentiality of such an achievement, but the inferences of such a potentiality in the characterisation of the situation' (1991: 192).

[47] Cf. Binnick (1991: 181), and also Pinkster (1990: 215).

(8)   Table II

situations/states of affairs

states
(non-dynamic)
'desire, want'

non-states
(dynamic)

activities
(atelic)
'run, swim'

performances
(telic)

achievements
(momentaneous)
'win the race'

accomplishments
(durative)
'run a mile'
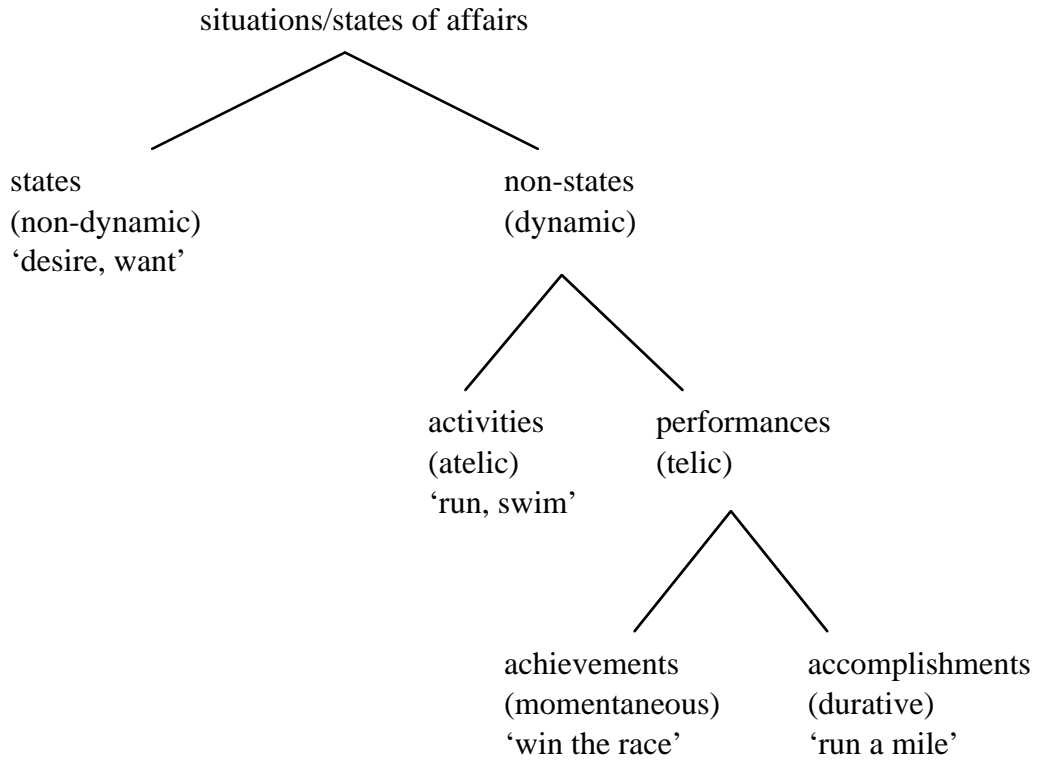
A slightly different interpretation of these categories is given by Moens and Steedman (1988). They attribute to non-states, which they label 'events', two features: atomic/extended and +/- consequent state. Including states, this gives five, not four, different types of situations: Vendler's class of activities is divided into two groups, points and processes. In the version of their table (1988: 17) reproduced below (Table III), I have added the traditional Vendlerian terms in brackets.

(9)   Table III

	EVENTS		STATES
	atomic (momentaneous)	extended (durative)	
+conseq (telic)	CULMINATION (achievements) recognize, spot, win the race	CULMINATED PROCESS (accomplishments) build a house, eat a sandwich	understand, love, know, resemble
-conseq (atelic)	POINT (activities) hiccup, tap, wink	PROCESS (activities) run, swim, walk, play the piano	

So far we have not considered repetition of situations. Freed (1979) accounts for these by dividing situations into 'single' and 'series'.[48] The former are classed as either durative, as in *he is sleeping*, or 'iterative', a different use of this term by which Freed denotes a set of uninterrupted subevents making up a whole event, as in, for example, *she is sneezing*. Series are repetitions of situations which may themselves be taken as single situations, and are similarly divided into two, serials and generics; all serials are said to be also generics, but generics are not all serials. Serials indicate that a series of self-contained events takes place over a period of time, to which specific reference is made; hence *he smokes a lot*. Generics do not make reference to repetition of a specific situation, and may hence be realized as regular repetitions or as general states: habitual occurrences (*he smokes a lot*), occupations (*she teaches college*), and general states (*they own a house*). Only imperfective situations are said to be able to occur in series, although in fact this generalization does not seem to be true.[49] Series may be formed from any other Vendlerian categories.

A different way of viewing series may be inferred from the work of Moens and Steedman (1988), who note that the addition of different features to different Vendlerian categories has various consistent effects; they produce a diagram of movements between the different areas of their table to map the various possibilities (1988: 18). They give the following example for the progressive state expressed by *Sandra is hiccupping*:

(10)   (point (*Sandra hiccup*))

(process (iteration (point (*Sandra hiccup*))))

(progressive (process (iteration (point (*Sandra hiccup*))))).

That is to say, the predicate *hicupped* (in *Sandra hiccupped (once)*) is to be classed as a point (activity), whereas *was hiccupping* (in *Sandra was hiccupping (all day long)*) is a progressive (a type of state), and the latter involves iteration of the former. In Moens and Steedman's terminology, the point has been 'coerced' into a state.

---

[48] Cf. Freed (1979: 32-6), summarized by Binnick (1991: 182-3) (although his page numbers differ).

[49] Cf. Hewson and Bubeník (1997: 15), where Hewson cites Russian imperfective *pit'* 'to drink', perfective *popit'* 'to have a drink', iterative *popivat'* 'to keep having drinks', and Dahl (1995: 419-20), who quotes an example from Serbo-Croatian but refers to the phenomenon as 'deviant behavior, since aspectual distinction[s]... tend to be neutralized in generics'. On the difference between forms in -σκον from present and aorist stems, see also note 20 above.

In this case, a series may therefore not be an individual verb-type, but merely the result of adding a feature to a situation — a feature we may variously call iteration, plurality of action or '+REPEATED'. We see in the above example what happens when this feature is added to the category point.

Much further work has been accomplished in recent years on the syntax and semantics of generic phenomena.[50] These relate to both the noun and to predicates, but the former do not concern Greek forms in -σκον. The relevant topic of verbs is more complicated, and involves the division of categories of types of sentences. The most fundamental distinction is between particular or episodic sentences, which denote a specific situation, and characterizing (i.e. generic) ones, which denote a property. A further division depends on whether the verb is dynamic or stative: it seems that, in the former case, a predicate may be used episodically and generically (the latter being termed the 'habitual' type), whereas any given stative predicate may be used either episodically or generically, but not as both (the latter being called the 'lexical' type of characterizing verb). Hence, in the table below, 'roar' is used in both episodic and characterizing sentences, but in the stative column there are two different predicates, 'weigh' and 'be'.[51] Hence:[52]

(11)   Table IV

	'Dynamic' situation	Stative situation
Episodic predicates	Episodic dynamic sentences *Simba roared*	Episodic stative sentences *Simba is (currently) in the cage*
Characterizing predicates	Habitual sentences *Simba roars when he smells food*	Lexical characterizing sentences *Simba (usually) weighs more than 200 lbs.*

The reason for the inverted commas around 'dynamic' is that, when such a verb is used in a habitual sentence, the sentence is deemed to be stative: it does not refer to a specific situation, but to a general property of the subject. In the characterizing predicate we are again dealing with the addition of repetition. We saw above:

(12)   point (*Sandra hiccupped*) + repetition   progressive state (*Sandra was hiccupping*).

---

[50] For a general introduction, see Krifka *et al.* (1995).

[51] For the criteria which lead to these divisions, see Krifka *et al.* (1995, esp. 1-18, 36-9).

[52] The table does not take account of the different types of NPs, but only of types of sentence; the total number of possible combinations is actually twelve (cf. Krifka *et al.* 1995: 18).

So here we can have the following:

(13) (a)   process (*Simba roars once*) + repetition
            habitual state (*Simba is always roaring*)

    (b)   process (*Simba roared*) + repetition
            habitual state (*Simba used to roar* [*whenever he smelled food*]).

By contrast, stative verbs can be either episodic or characterising, but unlike dynamic verbs they cannot be used in both types of sentence. So, *weigh* can only be used as a characterizing predicate:

(14) (a)   Episodic: **Simba is weighing 250 lbs. now
    (b)   Characterizing: Simba usually weighs more than 200 lbs.

and *be in the cage* is only episodic:

(15) (a)   Episodic: Simba is currently in the cage
    (b)   Characterizing: **Simba is being in the cage.

Interestingly, English 'used to' denotes a past characterizing event:

(16) (a)   'Dynamic': Simba used to roar when he smelled food (but he doesn't any more)
    (b)   Stative: Simba used to weigh more than 200 lbs. (but now he weighs less).

So, addition of 'used to' to a dynamic predicate gives a repeated series of events that make up a state, whereas addition to a stative results in a continuous state. This marker of habituality coerces different Vendlerian types with different effects, a vital concept for understanding Greek -σκον.

5. *Semantic analysis*

If we assume that -σκον is a marker of repetition, we should expect two different results from addition of this suffix to a stem; and this is exactly what we find. It has already been noted above that several scholars have labelled -σκον as 'intensive' on certain verbs (§1.2.2); it is now clear that this meaning merely results from a marker of iterativity being added to a stative predicate: from ναίω 'live' we expect continuous ναίεσκον 'used to live', the serial 'lived

repeatedly' being pragmatically unlikely.[53] Certain recent scholars have mentioned briefly that it is the meaning of the verb in question that is the important element here;[54] but in fact the best and most productive work that has been done in this field is also some of the very earliest work on -σκον — that of Stolpe (1849) and Týn (1859), and also Kluge (1911).

These scholars all classified iteratives in -σκον with an emphasis on syntactic and semantic criteria and, although their analyses are different, there are some consistently noted features. Stolpe (1849: 28-40) has four classes of forms, divided according to semantic and syntactic criteria; Týn (1859) deals first with forms derived from aorist stems and then those from presents; and Kluge (1911: 22, 26-30, 33-6), who had read Stolpe's work, discusses data from the *Iliad* and the *Odyssey* separately.

In what follows, forms in -σκον are divided into two classes. The first category corresponds to Freed's category of serials, and to Krifka *et al.*'s class of characterizing 'dynamics': these are instances where -σκον is added to a dynamic predicate, and involve an evidently repeated action. There are several sub-types within this class, and in these Stolpe's and Týn's detailed analyses will be followed. The second class comprises non-serials, or characterizing stative cases, where -σκον is found on a stative predicate.

5.1. Serial situations in Homer

Serial situations involve repetition of a dynamic situation, whether telic (achievements, accomplishments) or atelic (activities); once the situation has occurred, it happens again, although not necessarily immediately. Homer shows four different syntactic categories for serial situations. Of these four, the first three are easily classed, showing distinctive features which were noticed and discussed separately by Stolpe and Týn.[55]

5.1.1. 'Whenever X happened, then Y would happen'

Firstly, we find a major class of serials, where, when a situation occurs, one or more others follow, and this whole sequence of events is then repeated. In these cases, the earlier event is usually rendered by an optative in a subordinate clause,[56] and the others by forms in -σκον;[57] the form may therefore be

---

[53] Serial meanings are of course possible in context: cf. e.g. the examples of φιλέεσκον, both non-serial 'used to love' and serial 'used to have sex with, used to entertain' in §5.2 below.
[54] Cf. §6 below.
[55] Cf. Stolpe (1849: 36-8) and Týn (1859: 678-83 on aorist stems, 688-90 on present stems).
[56] This use of the optative is referred to by Týn as the 'Optativus frequentiae', and by Debrunner as the 'optativus iterativus' (Schwyzer and Debrunner 1939-71: 2.335-6); it was

summarized as 'whenever X happened, then Y would happen'. Stolpe (1849: 36) describes these instances as 'frequentativa'. Hence, from an aorist stem:

(17) ὅν τινα Τυδείδης ἄορι πλήξειε παραστάς,
        τὸν δ' Ὀδυσεὺς μετόπισθε λαβὼν ποδὸς ἐξερύσασκε, ... (*Il.* 10.489-90)
        But whoever the son of Tydeus stood beside and struck with his sword, him Odysseus, having seized by the foot from behind, would drag aside...

This type of construction is extremely frequent. Occasionally, the verb denoting the earlier action occurs in the simple imperfect, the participle or as a second iterative. In one very different instance we must understand the expression 'each time', when Odysseus fails three times to hug his mother:

(18) τρὶς μὲν ἐφορμήθην, ἑλέειν τέ με θυμὸς ἀνώγει,
        τρὶς δέ μοι ἐκ χειρῶν σκιῇ εἴκελον ἢ καὶ ὀνείρῳ
        ἔπτατ'· ἐμοὶ δ' ἄχος ὀξὺ γενέσκετο κηρόθι μᾶλλον, ... (*Od.* 11.206-8)
        Three times I rushed forward, and my heart bade me clasp her, and three times she flitted from my arms like a shadow or a dream; and [each time] there would come a sharp pain in the bottom of my heart.

5.1.2. 'At one time X would happen, at another time Y would happen'

Our second group is much smaller; here, indefinite adverbs of time are used to mark a repeated series of situations:[58] rather than the ordered sequence of 'whenever X happened, then Y would happen' we have the random 'at one time X would happen, at another time Y would happen'. In these cases, therefore, the connection between the situations concerned is generally looser, as:

(19) ἄλλοτε μέν τε Νότος Βορέῃ προβάλεσκε φέρεσθαι,
        ἄλλοτε δ' αὖτ' Εὖρος Ζεφύρῳ εἴξασκε διώκειν. (*Od.* 5.331-2)
        At one time the South Wind would throw it [the raft] forward to the North Wind to be carried along, and at another the East Wind would yield it to the West Wind to drive.

Sometimes the second situation of the pair is hard to understand as occurring without the first having also happened. In these cases, this second syntactic class

---

also noted by Delbrück (Brugmann and Delbrück 1893-1916: IV.63) and Wathelet (1973: 403).

[57] Cf. Stolpe (1849: 36-8), Týn (1859: 678-80, 688-9), and also Kluge (1911: 27, 33).

[58] Cf. Stolpe (1849: 38), Týn (1859: 680-2, 689), and also Kluge (1911: 26, 33).

seems close to the previous one: the difference is that equal emphasis is laid on both actions, rather than the earlier one being subordinated. As in the above example, both forms show -σκον. Hence, in the following description of ball-throwing, Homer says 'the one would throw, the other would catch' rather than 'when one had thrown, then the other would catch':

(20)   τὴν ἕτερος <u>ῥίπτασκε</u> ποτὶ νέφεα σκιόεντα
       ἰδνωθεὶς ὀπίσω· ὁ δ᾽ ἀπὸ χθονὸς ὑψόσ᾽ ἀερθεὶς
       ῥηϊδίως <u>μεθέλεσκε</u>, πάρος ποσὶν οὖδας ἱκέσθαι. (*Od*. 8.374-6)
       ... the one having bent back <u>would toss</u> it toward the shadowy clouds; and the other having leapt up from the ground <u>would</u> skilfully <u>catch</u> it in turn before his feet reached the ground.

5.1.3. 'One would do X'

In the third category, it is the subject, and not the time of the situation, which is varied:[59] the form therefore is 'someone would do X'. Almost all instances in Homer are accounted for by the twenty-eight occurrences of εἴπεσκε, as in:

(21)   ὧδε δέ τις εἴπεσκεν ἰδὼν ἐς πλησίον ἄλλον· ...
       ὣς φάσαν ἡ πληθύς· (*Il*. 2.271, 278)
       And thus one, having looked at his neighbour, <u>would say</u> [to him]:... So the multitude spoke.[60]

Apart from this, occurrences of this type are rare. An example of a present-stem form is found when Odysseus repeatedly binds together three rams as cover for his men's escape, and Homer says ὁ μὲν ἐν μέσῳ ἄνδρα φέρεσκε 'the one in the middle would bear a man' (*Od*. 9.429).

5.1.4. Other examples of repeated action

The above three groups are brought together by Stolpe and Týn, giving as they do instances of repeated action in context. This leaves one remaining group of serials to be discussed here, which indicate repeated action but not at any specific moment in time. The different treatments by Stolpe and Týn result in a different arrangement of remaining forms in -σκον for these scholars. Týn's analysis works on the basis of the distinction between iteratives from present stems and those from aorist ones: the above three types all occur for both tense

---

[59] Cf. Stolpe (1849: 38-9) and Týn (1859: 682-3, 689-90).
[60] Cf. the almost identical pattern at *Il*. 4.81, 4.85 and 22.372, 22.375.

stems, and this leaves him a group of forms in -σκον formed from present stems only, which he subdivides into two. One subset, made up of obviously serial formations, is our fourth syntactic class; in the other there is no interruption in the situation described, and these are in fact our stative forms (1859: 692-4; cf. §5.2 below). Stolpe, however, begins with what we call statives (1849: 30-3) and splits what comprises our fourth class, serials which indicate repeated action at an unspecified point in time, into those which largely indicate 'vel consuetudo ac mos vel munus et officium' (1849: 33; cf. Stolpe 1849: 33-6) and those where the action was repeated over some length of time, as indicated by the use of such adverbs as πολλά, πολλάκις, and αἰεί (1849: 39-40).

Following Stolpe and Týn, we therefore have left two sub-types of serials, which have in common the property that the situation concerned is repeated but not at a specific point in time. Examples of these two types of our fourth class are as follows. For that where there is no indicator such as πολλά, we for example hear of Thetis described as she

(22)   ἥ οἱ ἀπαγγέλλεσκε Διὸς μεγάλοιο νόημα. (*Il*. 17.409)
        who <u>would bring/used to bring</u> him [Achilles] news of the purpose of great Zeus.

For the other type, we have e.g.:

(23)   ὁ δ᾽ εὔκηλος μυχῷ Ἄργεος ἱπποβότοιο
        <u>πόλλ᾽</u> Ἀγαμεμνονέην ἄλοχον θέλγεσκεν ἐπέεσσιν. (*Od*. 3.263-4)
        but he [Aegisthus], at ease in a nook of Argos that is grazed by horses, <u>often</u> <u>would</u> <u>cajole/used to cajole</u> Agamemnon's wife with words.

5.1.5. Serial situations in Homer and Vendlerian categories

We may also note that all three categories of Vendlerian dynamic situations occur for serial events. For example, we have the momentaneous telic situation 'kill', an achievement, when Hermes says of Hector:

(24)   τὸν μὲν ἐγὼ μάλα πολλὰ μάχῃ ἔνι κυδιανείρῃ
        ὀφθαλμοῖσιν ὄπωπα, καὶ εὖτ᾽ ἐπὶ νηυσὶν ἐλάσσας
        Ἀργείους <u>κτείνεσκε</u>, δαΐζων ὀξέϊ χαλκῷ· (*Il*. 24.391-3)
        I have very often seen him with my eyes in battle that brings men honour, even when after driving the Argives to the ships he <u>used to slay</u> them, tearing [them] with the sharp bronze.

Similarly, a durative telic situation (accomplishment), when iterated in a generic fashion, reaches its culmination and is then repeated at a later time, as with πλΰνω 'wash (clothes)' when Homer speaks of the πλυνοί 'washing-troughs'

    (25)   ... ὅθι εἵματα σιγαλόεντα

          <u>πλύνεσκον</u> Τρώων ἄλοχοι καλαί τε θύγατρες (*Il.* 22.154-5)

          ... where the wives and fair daughters of the Trojans <u>used to wash</u> bright clothing.

With an atelic dynamic situation (activity), the action is understood as occurring at various points in time in the same way as a telic action is:

    (26)   τὸ δὲ <u>θαυμάζεσκον</u> ἅπαντες

          ὡς οἱ χρύσεοι ἐόντες ὁ μὲν λάε νεβρὸν ἀπάγχων, ... (*Od.* 19.228-9)

          And at this all men <u>used to marvel</u> how, though they were gold, the hound was strangling the fawn and pinning it...

The group of atelic situations where we find -σκον is especially notable since, although many forms in -σκον only occur once, we repeatedly find verbs of wearing (φορέεσκε occurs seven times; cf. also ζωννύσκετο at *Il.* 5.857); and of allowing (ἔασκον and εἴασκον are found four times each, nearly always with a negative).[61]

    In classifying forms in -σκον from aorist and present stems separately, Týn highlights the difference between serial and non-serial forms. A drawback to his approach, however, is that his breakdown only allows for forms in -σκον from aorist stems which fall into one of the first three categories given in this section. Yet in theory an aorist stem formation in this fourth category is possible, and in fact the one instance which Týn finds unclassifiable surely belongs here: it is αὐδήσασκε (cf. Týn 1859: 683). Hera shouts like Stentor,

    (27)   ὅς τόσον <u>αὐδήσασχ'</u> ὅσον ἄλλοι πεντήκοντα· (*Il.* 5.786)

          who <u>used to shout</u> as loudly as fifty other men...

This, however, is apparently the only such instance, unless — as is tempting — one classifies παρέβασκε (*Il.* 11.104) as the iterative of aorist ἔβην rather than as part of present βάσκω. In this case this ambiguous form would mean 'used to go beside'.[62] The lack of aorists in this category is surprising; presumably it is

---

[61] That is, seven times out of eight this verb is found with a negative, meaning 'not allow'; cf. Chantraine (1986-8: 1.319).

[62] Cf. note 18 above.

related to the semantics of the aorist itself, and its use for denoting single actions.[63]

5.2. Non-serial situations in Homer

We are left with the non-serials, the class of forms in -σκον which derive from verbs expressing states — that is, in our interpretation, the so-called 'intensives', where repeated action is not found. As noted above, Týn observed that these forms only occur on present stems; this is to be expected, as the situations they describe are by definition durative and atelic — we are not to understand that the state broke off and was repeated later, but that it was once constantly the case, and is so no longer. Obvious examples are the verbs of dwelling, ναιετάασκον (five times) and ναίεσκον (three), both 'dwelled, lived, used to dwell, live (in)'. Hence:

> (28)   Μαγνήτων δ᾽ ἦρχε Πρόθοος Τενθρηδόνος υἱός,
>   οἳ περὶ Πηνειὸν καὶ Πήλιον εἰνοσίφυλλον
>   <u>ναίεσκον</u>· (*Il*. 2.756-8)
>   And Prothous, son of Tenthredon, was leader of the Magnetes, who <u>lived/used to live</u>
>   around Peneius and Pelion, covered with trembling leaves.

Another frequent example is φιλέεσκον (eight times), which can usually be interpreted as a stative 'used to love', as in καί ἑ μάλιστα | δμῳάων φιλέεσκε (*Od*. 1.434-5) 'and she especially of all the handmaids used to love him [Telemachus]'. In two examples, however, this form appears to be used as an accomplishment 'have sex with',[64] and on one occasion the accomplishment 'entertain', when Diomedes kills Axylus, described as φίλος ... ἀνθρώποισι 'dear to men',

> (29)   πάντας γὰρ <u>φιλέεσκεν</u> ὁδῷ ἔπι οἰκία ναίων. (*Il*. 6.15)
>   for, living in a house by the road, he <u>used to give entertainment</u> to all.

Sometimes, both dynamic and stative interpretations are possible. So ἀριστεύεσκον (six times) is interpreted as stative 'used to be the best' (rather than accomplishment 'used to perform feats of valour') by Cunliffe (1924, s.v. ἀριστεύω), who lists all instances of the iterative under the meaning 'excel', literally (*Il*. 6.460, 11.746, 16.292, 551, 17.351) or metaphorically (*Il*. 11.627).

---

[63] On the difference between the meaning of the aorist and present forms, see note 20 above.
[64] *Il*. 9.450 and *Od*. 18.325; cf. Cunliffe (1924, s.v. φιλέω (2)).

Yet in the former set of cases it is ambiguous, as when Hector says to Andromache that, once he is dead, people will say of her:

(30) Ἕκτορος ἥδε γυνή, ὃς <u>ἀριστεύεσκε μάχεσθαι</u>
Τρώων ἱπποδάμων, ... (*Il.* 6.460-1)
There's the wife of Hector, who of all the horse-taming Trojans
<u>used to be pre-eminent in war/used to perform feats of valour</u>...

5.3. Iteratives in -σκον in Herodotus[65]

Since data from Herodotus are commonly used to argue that Homer's use of -σκον on -σα- aorists is a poetic extension, one may wonder how similar Herodotus' syntactic and semantic usage is. Here, serials of the type 'whenever X happened, then Y would happen' are again frequent, with the subordinate clause showing aorist optative[66] or a participle.[67] Hence, for example, Herodotus says that the Egyptian priests told him:

(31) ὡς ἐπὶ Μοίριος βασιλέος, <u>ὅκως ἔλθοι</u> ὁ ποταμὸς ἐπὶ ὀκτὼ πήχεας τὸ ἐλάχιστον, <u>ἄρδεσκε</u> Αἴγυπτον τὴν ἔνερθε Μέμφιος. (Hdt. 2.13.1)
when Moeris was king, <u>whenever</u> the river <u>rose</u> as least eight cubits, it <u>would water</u> [all] Egypt below Memphis.

There are also occasional examples of the second class of serials, the unordered 'at one time X would happen, at another time Y would happen', found in Herodotus' explanation of how the people of Barce found the besieging Persians' mines:[68]

(32) τὰ μέν νυν ὀρύγματα ἀνὴρ χαλκεὺς ἀνεῦρε ἐπιχάλκῳ ἀσπίδι ὧδε ἐπιφρασθείς· περιφέρων αὐτὴν ἐντὸς τοῦ τείχεος προσῖσχε πρὸς τὸ δάπεδον τῆς πόλιος· <u>τὰ μὲν δὴ ἄλλα ἔσκε</u> κωφά, πρὸς ἃ προσῖσχε, κατὰ <u>δὲ</u> τὰ ὀρυσσόμενα <u>ἠχέεσκε</u> ὁ χαλκὸς τῆς ἀσπίδος. (Hdt. 4.200.2)

---

[65] The text used here is that of Rosén (1987-97).

[66] Cf. also Hdt. 1.186 ἀπαιρέεσκον, ἐπιτείνεσκε; 1.196 ἐσάγεσκον, πωλέεσκε; 2.13 ἄρδεσκε; 2.174 ἄγεσκον, κλέπτεσκε; 4.42 σπείρεσκον, μένεσκον; 4.43 φεύγεσκον; 4.78 λάβεσκε, κατελίπεσκε; 4.129 ἔσκον; 4.130 ἐλάβεσκον; 6.12 ἔχεσκε; 7.41 μετεκβαίνεσκε; 7.119 σιτεύεσκον, ἔσκε (once or twice), ποιεέσκετο, ἔχεσκον, ἀπελαύνεσκον; 7.211 φεύγεσκον; and 9.74 βαλέσκετο.

[67] Cf. Hdt. 1.36 διαφθείρεσκε, ποιέεσκον; 1.100.1 ἐσπέμπεσκον, ἐκπ έμπεσκε; 1.148 ἄγεσκον; 3.117 ἄρδεσκε; 3.119 κλαίεσκεν, ὀδυρέσκετο; 7.33 ἔρδεσκε; and 9.40 ἔσκον.

[68] Cf. also Hdt. 1.196 ἔσκον, ἔσκε; and 4.128 τρέπεσκε.

As for the mines, a smith discovered them by means of a bronze shield, having devised thus: carrying it around within the wall he held it against the ground of the city; <u>some places</u> at which he held it <u>would be</u> [would sound] dull, <u>but</u> along the mines the bronze of the shield <u>would ring</u>.

There are no examples of the Homeric type 'someone would do X'. Serials do also however occasionally occur in Herodotus without limiting context, as when Mardonius tries to persuade the new king to march against Athens next:[69]

(33) οὗτος μέν οἱ ὁ λόγος ἦν τιμωρός, τούτου δὲ τοῦ λόγου παρενθήκην <u>ποιεέσκετο</u> τήνδε, ὡς ἡ Εὐρώπη περικαλλὴς χώρη καὶ δένδρεα παντοῖα φέρει τὰ ἥμερα ἀρετήν τε ἄκρη βασιλέϊ τε μούνῳ θνητῶν ἀξίη ἐκτῆσθαι. (Hdt. 7.5.3)

This was an argument for vengeance; but he <u>would</u> [repeatedly] <u>make</u> an addition of the argument, that Europe was a very beautiful country, and it bore all kinds of cultivated trees, a land high in excellence, worthy to be owned by the king alone of mortals.

Non-serials in -σκον may be non-existent in Herodotus: there are only two possible examples. Firstly, there is one occasion when ἔσκον may be translated as stative 'used to be'; one may also read serial sense here, however, in its use with ἑκάστοτε 'each time, on each occasion':

(34) οἱ δὲ Πάριοι ὅκως μέν τι δώσουσι Μιλτιάδῃ ἀργύριον, οὐδὲ διενοεῦντο, οἳ δὲ ὅκως διαφυλάξωσι τὴν πόλιν, τοῦτο ἐμηχανῶντο ἄλλα τε ἐπιφρασσόμενοι, καὶ τῇ μάλιστα <u>ἔσκε</u> ἑκάστοτε ἐπίμαχον τοῦ τείχεος· τοῦτο ἅμα νυκτὶ ἐξήρετο διπλήσιον τοῦ ἀρχαίου. (Hdt. 6.133.3)

And the Parians did not at all intend to give money to Miltiades, but to guard their city carefully, and they contrived this and other things, and blocking up that part of the wall where on each occasion it <u>used to be</u> vulnerable; while it was night, it was raised to double its old [height].

Secondly, we may also understand ἐθέλεσκον as stative in the following passage:

(35) ταῦτα ἔλεξαν, καὶ μετὰ ταῦτα αὐτίκα πείθεσθαι οὐδεὶς ἤθελε, ἀλλά, οἷα στρατιή, σκηνάς τε πηξάμενοι ἐν τῇ νήσῳ ἐσκιητροφέοντο καὶ ἐσβαίνειν οὐκ <u>ἐθέλεσκον</u> ἐς τὰς νέας οὐδ' ἀναπειρᾶσθαι. (Hdt. 6.12.4)

---

[69] Cf. also Hdt. 4.78 ποιέεσκε; 7.5 ποιεέσκετο; and 7.106 πέμπεσκε.

> They [the Ionians] said these things, and after this no man was willing to obey; but, having pitched themselves tents on the island, as though they were an army, they kept from the sun, and <u>were not willing</u> to embark in their ships, nor to exercise them.

On the other hand, ἐθέλεσκον could denote a serial situation: the suffix -σκον could be being used to indicate that the men are repeatedly being asked to board their ships, and are repeatedly refusing.

Occasionally, we find ἄν used with iteratives in Herodotus (2.174, 3.119, 4.42, 4.78, 4.130), presumably a development which reflects the classical usage of simple aorists and imperfects with ἄν instead of iterative forms in -σκον.[70]

Herodotus, therefore, uses forms in -σκον perhaps entirely with serial meaning, and especially with the aorist optative; entirely absent is the type 'someone would do X', and also expressions of the atelic activities type 'used to wear', quite common in Homer (cf. §5.1.5 above). This could be because Herodotus liked to use -σκον to mark serial situations, of which 'when X happened, then Y would happen' is the most frequent type both in his work and in epic. Alternatively, it could indicate that, as is the case with forms in -σκον derived from -σα- aorists, non-serial use of -σκον is a poetic innovation on the part of Homer.

6. Conclusion: the semantics of forms in -σκον

It is now clear that stative forms in -σκον have given rise to much confusion and often prevented the formations in -σκον from being seen as a coherent whole. Stolpe (1849) argued that both presents in -σκω and preterites in -σκον showed iterative meaning; although this view does not seem acceptable, his argument for coherence within the latter group is now backed up by modern theory. Certain scholars have recently mentioned that differences arise from the meanings of the verbs involved,[71] but none of them attempts to classify the verbs semantically.

Is our suffix therefore to be described as iterative, durative, or iterative-durative? It primarily seems to denote iteration of the action concerned, which is realized as repeated (serial) in dynamic situations and usually as continuous (non-serial) in stative ones. Indeed, the latter use is so rare in Herodotus that the

---

[70] For this use cf. Ruijgh (1985: 147) and Rijksbaron (1994: 14-15).

[71] Cf. Wathelet (1973: 403), Puhvel (1991: 13), and Sihler (1995: 506); for example, Puhvel contrasts durative 'kept smiling until the face muscles got tired' with iterative 'kept smiling in various situations'.

former would seem to have been regarded as fundamental, in the Classical era at least.

Data from other languages parallel the Greek usage. English *used to*, which also indicates repeated/continuous action in the past, cannot be used to denote serial situations which occur in a limited time-frame, as -σκον frequently does in both Homer and Herodotus: hence we would translate (18) above (*Od.* 11.206-8) as 'Three times I rushed forward, ... and [each time] there would come/**used to come a sharp pain'.[72] A suffix with a wider range still than Greek -σκον is West Greenlandic Eskimo *-sar-/-tar-*.[73] This suffix, unlike -σκον, is used in the present:

(36) (a) *allakkanik allapoq* 'he is writing/was writing/wrote letters'
(b) *allakkanik allattarpoq* 'he (usually) writes letters'
(c) *arfinermut makittarpunga* 'I get up at six o'clock'
(d) *miaartortarput* 'they (i.e. cats) meow'.

It is also used, like -σκον, to mark the second event in serial situations of the type 'when X happened, then Y would happen':

(37) *qitsuk qinngasaarnoqaraangami miaartortarpoq* 'if you tease a cat, it meows'.

In addition to these data, also of interest from Dahl's (1995) typological account is the statement that generics tend to be 'minimally marked for tense-aspect'.[74] Is the absence of the augment therefore to be connected with the characterizing properties of the suffix?

If the preceding account is correct, therefore, the function of Greek -σκον is no longer as isolated as it once seemed; though long regarded as an unusual formation within Greek, it clearly fits well within a much wider picture.

References

For works of classical authors and epigraphical publications, the same abbreviations and editions are used as are found in LSJ (xvi-xlv, supplement x-xxxi).

Adams, Douglas Q. (1988). *Tocharian Historical Phonology and Morphology* (American Oriental Series, 71). New Haven.

---

[72] On English *used to*, see further Quirk *et al.* (1985: 140).
[73] Data are taken from Dahl (1995: 422-3); cf. also Dahl (1985: 100-2).
[74] Dahl (1995: 415); see further Dahl (1995: 415-16, concluded at 425).

Allen, Andrew Strachan (1980). *The Development of the Inchoative Suffix in Latin and Romance*. Ph.D. thesis, University of California, Berkeley.

Allen, Thomas W. (ed.) (1917-19). *Homeri Opera: Tomi III-IV* (2nd edn.). Oxford.

Bechtel, Friedrich (1921-4). *Die griechischen Dialekte*. Berlin.

Bechtel, George (1936). *Hittite Verbs in -sk-: A Study of Verbal Aspect*. Ann Arbor.

Binnick, Robert I. (1991). *Time and the Verb: A Guide to Tense and Aspect*. New York/Oxford.

Bottin, Luigi (1969). 'Studio dell'aumento in Omero'. *SMEA*, 10: 69-145.

Brugmann, Karl (1902). 'Die ionischen Iterativpräterita auf -σκον'. *IF*, 13: 267-77.

—, and Delbrück, Berthold (1893-1916). *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen* (2nd edn.). Strassburg.

Bybee, Joan. L. (1992). 'Tense, aspect and mood', in William Bright (ed.), *International Encyclopedia of Linguistics*. New York/Oxford, 4.144-5.

Chantraine, Pierre (1986-8). *Grammaire homérique* (6th edn.). Paris.

— (1991). *Morphologie historique du grec* (3rd edn.). Paris.

Clackson, James (1994). *The Linguistic Relationship between Armenian and Greek* (Publications of the Philological Society, 30). Oxford.

Comrie, Bernard (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge.

Cunliffe, Richard J. (1924). *A Lexicon of the Homeric Dialect*. Glasgow.

Curtius, Georg (1852). 'Vermischte etymologien'. *KZ*, 1: 25-36.

— (1877-80). *Das Verbum der griechischen Sprache* (2nd edn.). Leipzig.

Dahl, Östen (1985). *Tense and Aspect Systems*. Oxford.

— (1995). 'The marking of the episodic/generic distinction in tense-aspect systems', in Greg N. Carlson and Francis Jeffry Pelletier (eds.), *The Generic Book*. Chicago-London, 412-25.

Dressler, Wolfgang (1968). *Studien zur verbalen Pluralität. Iterativum, Distributivum, Durativum, Intensivum in der allgemeinen Grammatik, im Lateinischen und Hethitischen* (Österreichische Akademie der Wissenschaften, Philosophisch-Historische Klasse, Sitzungsberichte, 259). Vienna.

Duhoux, Yves (1992). *Le verbe grec ancien: Éléments de morphologie et de syntaxe historiques*. Louvain-la-Neuve.

Ernout, Alfred (1953). *Morphologie historique du latin* (3rd edn.). Paris.

Fantini, Julio (1950). 'Vista de conjunto sobre los pretéritos iterativos jónicos en -σκον'. *Helmantica*, 1: 319-38.

Freed, Alice F. (1979). *The Semantics of English Aspectual Complementation.* Ph.D. thesis, University of Pennsylvania (1976).

Gehring, August (1891). *Index Homericus.* Leipzig.

Giacalone Ramat, Anna (1967). 'La funzione del suffiso -σκ- nel sistema verbale greco'. *AGI*, 52: 105-23.

Hackstein, Olav (1995). *Untersuchungen zu den sigmatischen Präsensstamm-bildungen des Tocharischen* (*HS* Ergänzungsheft, 38). Göttingen.

Haverling, Gerd (2000). *On Sco-Verbs, Prefixes and Semantic Functions: A Study in the Development of Prefixes and Unprefixed Verbs from Early to Late Latin* (Studia Graeca et Latina Gothoburgensia, 64). Göteborg.

Hewson, John, and Bubeník, Vít (1997). *Tense and Aspect in Indo-European Languages: Theory, Typology, Diachrony* (Current Issues in Linguistic Theory, 145). Amsterdam/Philadelphia.

Keller, Madeleine (1992). *Les verbes latins à infectum en* -sc-*, étude morphologique à partir des formations attestées dès l'époque préclassique* (Collection Latomus, 216). Brussels.

Kimball, Sara E. (1980). 'A Homeric note'. *Glotta*, 58: 44-6.

Klingenschmitt, Gert (1982). *Das altarmenische Verbum.* Wiesbaden.

Kluge, Heinrich (1911). *Syntaxis Graecae quaestiones selectae.* Diss. Berlin.

Krifka, Manfred, Pelletier, Francis Jeffry, Carlson, Greg N., ter Meulen, Alice, Chierchia, Gennaro, and Link, Godehard (1995). 'Genericity: an introduction', in Greg N. Carlson and Francis Jeffry Pelletier (eds.), *The Generic Book.* Chicago/London, 1-124.

Kronasser, Heinz, and Neu, Erich (1966-87). *Etymologie der Hethitischen Sprache.* Wiesbaden.

Kujọṛẹ, Ọbafẹmi (1973). *Greek Polymorphic Presents: a Study of their Development and Functional Tendencies.* Amsterdam.

Leumann, Manu, Hofmann, J.B., and Szantyr, Anton (1965-79). *Lateinische Grammatik.* Munich.

Liddell, Henry G., Scott, Robert J., Jones, Henry S., and McKenzie, Roderick (1996). *A Greek-English Lexicon, with a Revised Supplement* (9th edn.). Oxford.

LSJ: see Liddell, Scott, Jones, and McKenzie (1996).

Meier-Brügger, Michael (1992). *Griechische Sprachwissenschaft.* Berlin/New York.

Meillet, Antoine, and Vendryes, Joseph (1948). *Traité de grammaire comparée des langues classiques* (2nd edn.). Paris.

Moens, Marc, and Steedman, Mark (1987). 'Temporal ontology and temporal reference'. *Computational Linguistics*, 14: 15-28.

Monro, David B., and Allen, Thomas W. (eds.) (1920). *Homeri Opera: Tomi I-II* (3rd edn.). Oxford.

Negri, Mario (1976). 'Studi sul verbo greco II'. *Acme*, 29: 233-43.

Oettinger, Norbert (1979). *Die Stammbildung des hethitischen Verbums* (Erlanger Beiträge zur Sprach- und Kunstwissenschaft, 64). Nürnberg.

Pinault, Georges (1989). 'Introduction au tokharien'. *LALIES: Actes des sessions de linguistique et de littérature*, 7: 3-224.

Pinkster, Harm (1990). *Latin Syntax and Semantics* (revised translation of *Lateinische Syntax und Semantik*, by H. Mulder). London/New York.

Porzig, Walter (1927). 'Zur Aktionsart indogermanischer Präsensbildungen'. *IF*, 45: 152-67.

Puhvel, Jaan (1991). *Homer and Hittite* (IBS Vorträge und kleinere Schriften, 47). Innsbruck.

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey, and Svartvik, Jan (1985). *A Comprehensive Grammar of the English Language*. London.

Rijksbaron, Albert (1994). *The Syntax and Semantics of the Verb in Classical Greek: An Introduction* (2nd edn.). Amsterdam.

Risch, Ernst (1974). *Wortbildung der homerischen Sprache* (2nd edn.). Berlin/New York.

Rix, Helmut (1986). *Zur Entstehung des urindogermanischen Modussystems* (IBS Vorträge und kleinere Schriften, 36). Innsbruck.

— (1992). *Historische Grammatik des Griechischen: Laut- und Formenlehre* (2nd edn.). Darmstadt.

— (ed.) (1998). *Lexikon der indogermanischen Verben: Die Wurzeln und ihre Primärstammbildungen* (unter Leitung von H. Rix und der Mitarbeit vieler anderer bearbeitet von M. Kümmel, T. Zehnder, R. Lipp und B. Schirmer). Wiesbaden.

Rosén, Haiim B. (1962). *Eine Laut- und Formenlehre der herodotischen Sprachform*. Heidelberg.

— (ed.) (1987-97). *Herodoti Historiae*. Stuttgart/Leipzig.

Ruijgh, C. J. (1985). 'Problèmes de philologie mycénienne'. *Minos*, 19: 105-67 (= *Scripta Minora ad linguam graecam pertinentia*. Amsterdam, 1991-6, 2.43-105).

Ruipérez, Martín S. (1954). *Estructura del sistema de aspectos y tiempos del verbo griego antiguo: análisis funcional sincrónico* (Theses et studia philologica Salmanticensia, 7). Salamanca.

Schwyzer, Eduard, and Debrunner, Albert (1939-71). *Griechische Grammatik auf der Grundlage von Karl Brugmanns Griechischer Grammatik*. Munich.

Shipp, G. P. (1972). *Studies in the Language of Homer* (2nd edn.). Cambridge.

Sihler, Andrew L. (1995). *New Comparative Grammar of Greek and Latin*. New York/Oxford.

Stolpe, August (1849). *Iterativorum graecorum vis ac natura ex usu Homeri atque Herodoti demonstrata*. Diss. Bratislava.

Szemerényi, Oswald J. L. (1996). *Introduction to Indo-European Linguistics* (translated from *Einführung in die vergleichende Sprachwissenschaft*, 4th edn., 1990, with additional notes and references). Oxford.

Týn, Emmanuel (1859). 'Über den Gebrauch und die Bedeutung der iterativen Imperfecta und Aoriste im Griechischen'. *Zeitschrift für die österreichischen Gymnasien*, 10: 677-95.

Vine, Brent (1993). 'Greek -ίσκω and Indo-European "*-isk̑e/o-"'. *HS*, 106: 49-60.

Wathelet, Paul (1973). 'Études de linguistique homérique'. *AC*, 42: 379-405.

Watkins, Calvert. (1969). *Indogermanische Grammatik: Band III/1: Formen-lehre [Geschichte der Indogermanischen Verbalflexion]*. Heidelberg.

Zerdin, Jason (1998). 'The semantics of verbs in -σκω in Plato: a study of minimal pairs'. *Oxford University Working Papers in Linguistics, Philology and Phonetics*, 3: 125-154.

— (2000). *Studies in the Ancient Greek Verbs in* -skō. D. Phil. thesis, Oxford.

— (forthcoming). 'The semantics of verbs in -σκω in Plato: a study of minimal pairs'. *Mnemosyne*.

Notes on some Sabellic demonstratives

J. H. W. Penney

1. The Latin anaphoric pronoun *is, ea, id* shows an alternation between a zero-grade stem *i-*, seen in nom. sg. m. *is*, nom.-acc. sg. n. *id*, and a full-grade thematic stem in **eyo-* (f. **eyā-*), seen in acc. sg. m. *eum*, abl. sg. m. and n. *eō*, nom. sg. f. *ea*, acc. sg. f. *eam*, abl. sg. f. *eā*, etc.); the gen. sg. form for all genders is *eius*, which is widely believed to continue a redetermined form of **esyo*, with a simple stem **e-*, on the strength of a comparison with Skt. *ásya* (gen. sg. m. and n.).[1]

The Sabellic pronoun that corresponds to this both in form and in function shows a similar alternation between **i-* and **eyo-* in the forms of the nominative and accusative, where the pattern is clearly comparable to that of Latin. Elsewhere the Sabellic pronoun is built on a stem **eiso-* (f. **eisā-*) that is not familiar from Latin; it has been variously explained, most plausibly perhaps as arising by reanalysis of an inherited gen. pl. m. and n. **ei-sōm* (cf. Skt. *eṣā́m*) with pronominal ending **-sōm* as **eiso-* + ending **-ōm*, but whatever its origin it is characteristic of all cases other than the nominative and accusative. Not all case-forms occur in the texts, but the following may be cited for illustration, supplemented by one or two forms of 'the same' in Umbrian, formed from this anaphoric pronoun with a suffixed particle *–ont*.[2] (On U. **esmik**, *esmei*, which some would include within this paradigm, see §11.)

	Oscan			Umbrian		
sg.	m.	f.	n.	m.	f.	n.
nom.	*izic*	**íúk**	**ídík**	**ere(k)**		**eřek**
acc.	*ionc*	**íak**	**ídík**		*eam*	**eřek**
gen.	**eíseís**			*erer*	*erar*	
abl.	**eísúd**	**eísak**		**eru-ku**	**erak**	
loc.	**eíseí**	**e]ísaí**				

---

[1] See Leumann (1977: 466-7), Meiser (1998: 159-60).

[2] For the paradigms see Buck (1928: 141); for a full list of attestations see Untermann (2000: 229-30, 355-60). For an understanding of the Umbrian forms, it may be helpful to know that Umbrian shows rhotacism of intervocalic *s* and, in its later forms, final -*s*, that the diphthong **ei* is monophthongized to a long vowel written <e>, and that original short **i* gave a vowel that may be written either <i> or <e>.

	Oscan			Umbrian		
pl.	m.	f.	n.	m.	f.	n.
nom.	*iusc*			*eur-ont*		
acc.			*ioc*		**eaf**	*eo*
gen.	**eisunk**	*eizazunc*		**eru**		
dat.-abl.	*eizois*	*eiza(i)sc*			**ererunt**	*erir-ont*

2. This division of the paradigm, between nominative and accusative forms on the one hand and a sigmatic extended stem for the other cases, is also found with the Oscan pronoun *eko-/ekso-* 'this', equivalent in function to Latin *hic*:[3]

sg.				pl.			
	m.	f.	n.		m.	f.	n.
nom.		**ek(úk)**				**ekas**	
acc.	**ekúk**	**ekak**	**ekík**			**ekass**	
abl.	**eksuk**	*exac*				*exaisc-en*	
loc.	*exeic*						

In addition there is an adverb **ekss**, *ex* meaning 'thus', which must somehow be formed on the sigmatic stem although the details are obscure.

Comparable forms may be cited from Paelignian and Marrucinian: nom. sg. f. *ecuc* (Ve. 213/Pg 9); acc. sg. f. *ecan* (Po. 206/MV 3); nom.-acc. sg. n. *ecic* (Ve. 213/Pg 9); but gen. pl. m. *esuc* (Ve. 218/MV 1). One may also note the Paelignian adverb *ecuf* 'here' (Ve. 214/Pg 10), formed from the non-sigmatic stem.

3. The origin of this paradigm is far from clear: both **e-* and **-k-* are well-known deictic elements so that a ready etymology might thus be found for the basic stem **eko-*, but there is no agreement as to whether the sigmatic forms arise from the addition of a pronoun **so-* or through a process of reanalysis of the genitive plural similar to that posited by some for the anaphoric pronoun (see above), or quite simply by analogy with the alternation in the forms of this pronoun. Whatever the explanation, the parallel between the two pronominal paradigms is striking and can hardly be accidental.

4. In Umbrian there is a pronominal stem *es(s)o-* 'this'. Forms attested in the Iguvine Tables are: gen. pl. n. *esom-e*; abl. sg. m./n. **esu-ku**, *es(s)u*; abl. sg. f.

---

[3] For the attested forms see Untermann (2000: 216-17).

esa; abl. pl. n. *esis-co*, *esir*. There is also an adverbial form **esu(k)**, *eso(c)* 'thus, as follows' from the same stem.

Forms of this pronoun occur in prayers accompanying sacrificial offerings (in twenty instances), e.g. TI VIa 25 *di grabouie tio esu bue peracrei pihaclu ...* 'Jupiter Grabovius, (I invoke) you with this *p.* ox as an expiatory offering ...' In this context, with reference to the animal or object being presented, Latin would use forms of *hic*, and a similar value 'this' may plausibly be assigned to the Umbrian pronoun. Elsewhere the pronoun is used in reference to the ceremonies actually being performed (TI Ib 8, IV 29-30, VIb 47), and it occurs twice in a phrase spoken during the lustration of the *poplo-* (the citizen body *qua* army) requiring specified foreigners to depart *ehesu poplu/ehe esu poplu* 'from this *poplo-*'. In all these cases the reference is to something in the immediate context, as has been universally recognized.

The adverb **esu(k)** is used in the Tables in the majority of instances to introduce the actual words to be spoken during the various rituals, and twice to present the text of resolutions passed by the brotherhood of priests (Va 1, 14); it has a similar function in VIa 8, where the augural *templum*, we are told, *eso tuderato est* 'is bounded as follows', and a list of the boundaries is duly given. There are two instances (VIa 20, VIIb 3) of a form *iso*, *issoc* being used with a following clause introduced by *pusi/pusei* to mean something like 'in such a way that, in such a way as', e.g. VIIb 3 *sue neip portust issoc pusei subra screhto est ...* 'if he shall not have brought (them) in such manner as is written above ...', and it is not entirely clear if the spelling difference with *i-* in precisely these two instances should be taken to mean that this is to be treated as a different adverb (see Untermann 2000: 348-9); it is certainly true that the function of the adverb in these two instances is more anaphoric than presentational.

5. The function of these forms seems perfectly straightforward, but a complication arises from the fact that in Umbrian there is another pronoun with an apparently rather similar function, namely *esto-*, which raises the question of the deictic system as a whole. If it seems awkward to assume that Umbrian had two demonstrative stems with identical function, it is understandable that scholars should have tried, albeit at times half-heartedly, to find some point of difference between them. The stem *es(s)o-* may correspond functionally to Latin *hic* not simply in terms of near reference but more specifically with first-person deixis, so that *esto-* could be taken to correspond functionally (as also no doubt etymologically) to Latin *iste*, with second-person deixis (so explicitly Muller 1926: 222).

This interpretation often emerges only from translations of the Umbrian terms into Latin with *es(s)o-* rendered by *hic* and *esto-* by *iste* (so, for instance,

Conway 1897, Vetter 1953, Ernout 1961, Meiser 1998); in such cases it is not always clear that the scholars are committed to a difference of function rather than trying to mirror in their translations the different stems used in Umbrian. Elsewhere it appears that uncertainty prevails: von Planta (1892-7) gives *hic* as the translation of *es(s)o-* in his glossary, but for *esto-* he gives both *hic* and *iste* (yet maintaining consistent translation by *iste* in his Latin version of the texts of the Tables); Buck (1928: 142-3) states that *es(s)o-* corresponds in use to *hic*, but *esto-* is simply translated as *iste* without any statement about its use; Devoto (1963) gives the usual glosses *hic* and *iste* in his *Index Verborum*, but in his translation consistently uses *hic* for both. Poultney (1959), on the other hand, translates both sets of forms into English with 'this', and in his discussion of demonstratives states (110) that *es(s)o-* has the same semantic value as Latin *hic*, while *esto-* is used sometimes with reference to what has just been mentioned, sometimes to that which is newly introduced, but it 'has not the special nuance of association with the second person'. More recently Untermann (2000: 236) notes of *esto-* that it is 'auf Aktuelles, unmittelbar Bevorstehendes oder gegenwärtig Gültiges verweisend', and is equivalent in function to Latin *hic* rather than *iste*. The instances from the Tables certainly bear this out, and show that the uses of *es(s)o-* and *esto-* are essentially identical.

6. There are eleven instances of *esto-* in the Iguvine Tables. One of them (IIb 24 **iupater saçe tefe estu vitlu vufru sestu** 'Jupiter *S.*, to you I present this *v.* calf') clearly refers to the sacrificial animal being offered at the time of speaking; four (Ia 1 = VIa 1, IIa 2, VIa 56) refer to the ceremonies actually in progress or being initiated. Exactly similar functions were performed by *es(s)o-* (see §4). There are three instances of reference to something that has just been said: VIb 62, 63 *ape este dersicust* 'when he shall have said this', VIIa 51 *este trioper deitu* 'he is to say this three times' (VIb 62, 63), to which one might add the two instances of reference back to the list of boundaries given immediately before: VIa 15 *hondra esto tuderor porsei subra screihtor sent* 'below these boundaries which are written above', and VI1 15-16 *subra esto tudero* 'above these boundaries'. Finally there is an instance of reference forward, introducing the actual words that follow directly: IIb 23 **estu iuku habetu** 'use these words', which is functionally similar to the use of adverbial **esu(k)** noted in §4.

7. The overlap of functions between *es(s)o-* and *esto-* in the Iguvine Tables would seem evident. In the scanty other Umbrian inscriptions there are only a few attestations of either stem: in Ve. 233/Um 8 a nom. sg. f. *eso* occurs in the sentence *cubrar matrer eso beo* 'this fountain belongs to (lit. 'is of') *C.* Mater (a deity)', in an inscription on a stone cistern, with reference to the object itself; in

Po. 7/Um 5 a nom.-acc. pl. n. **estac** appears in agreement with **vera** 'gate' in a badly broken building inscription, presumably with reference to the structure of which the stone originally formed part. Both of these would most naturally be rendered in Latin with forms of *hic*.

There are also early attestations of the pronoun *esto-* outside Umbria, in certain closely related Sabellic languages. In a 'Pre-Samnite' inscription on a bronze stamnos from Campania (Ve. 101/PS 3) we read: *vinuχs veneliis peracis estam tetet venilei viniciiu* 'V., son of V., P. gave this to V., son of V.', where *estam* must refer to the object itself. In a South Picene inscription (AP 3) there is a form **estas**, possibly nom. pl. f., but the context is too broken for any certainty over the syntax or the interpretation of the reference. There are also two instances in South Picene of adverbial **estuf(k)** (AP 2, TE 5), most plausibly in the context of a funerary inscription to be rendered 'here', with which one may compare Pael. *ecuf* in *ecuf incubat* 'lies here' (Ve. 214/Pg 10), although the two South Picene inscriptions still resist complete interpretation.

None of these occurrences of *esto-* (and none of those in the Iguvine Tables) seems to require interpretation as having a function different from that of *es(s)o-*, yet various Italian scholars have in recent years claimed explicitly that *esto-* does in fact have second-person deixis, and that in Umbrian, 'Pre-Samnite', and South Picene there is a tripartite system of deixis just like that of Latin.

8. Prosdocimi (1978: 742-3, 1979: 168), starting from the need to distinguish functions for the different demonstrative stems in the Iguvine Tables, claims that *esto-* is the deictic corresponding both functionally and etymologically to Lat. *iste*. The second person implied in this context is the user of the Tables, the one who will read and follow the instructions there laid down. He notices that there will be a problem at TI IIb 23-4 **iupater saçe tefe estu vitlu vufru sestu** 'Jupiter S., to you I present this *v.* calf', which is not part of the instructions but comes within a prayer addressed to a deity, so that the implicit second person here cannot be the reader; he is also troubled by the fact that at every other presentation of a sacrificial victim or offering the pronoun *es(s)o-* is used (e.g. VIa 25 *di grabouie tio esu bue ...* 'Jupiter Grabovius, (I invoke) you with this ox ...'). Prosdocimi's solution to these problems is to argue that this exceptional use of **estu** is due to rhythmic factors, in effect the rhyme with **sestu**, citing as a parallel Cato *de agr.* 134 *Iuppiter, macte isto ferto esto* 'Jupiter, be honoured with that *fertum*', where *isto* echoes *esto* 'be', beside the usual formulaic *hic* for offerings in such phrases such as *hac strue*, *hoc ferto*, etc.[4]

---

[4] L. *fertum* is a kind of sacrificial cake, *strues* something similar.

The Cato passage is interesting as providing the only instance of *iste* in prayers in Cato, and it is worth looking more closely at its context. Prayers attending the sacrifice of a sow (*porca praecidanea*) before the harvest are addressed here to both Janus and Jupiter, but two prayers to each, the first accompanying solid offerings (a *strues* and a *fertum* respectively), the second offerings of wine (*uinum inferium*). In the first prayers we find, as expected, forms of *hic*: *Iane pater, te hac strue ommovenda bonas preces precor ...* 'Father Janus, in offering this *strues* I pray good prayers to you ...', *Iupiter, te hoc ferto obmovendo bonas preces precor ...* 'Jupiter, in offering this *fertum* I pray good prayers to you'. The second prayer to Janus runs: *Iane pater, uti te strue ommovenda bonas preces precatus sum, eiusdem rei ergo macte uino inferio esto* 'Father Janus, just as in offering the *strues* I prayed good prayers to you, for that same reason be honoured with an offering of wine', where there is reference in the past tense to the offering of the *strues* (with no demonstrative pronoun) but the presentation this time is of wine. The parallel second prayer to Jupiter is similarly structured but more succinct: *Iupiter, macte isto ferto esto, macte uino inferio esto.* 'Jupiter, be honoured with that *fertum*, be honoured with an offering of wine'. Once again there is a reference back to the solid offering, but it is not being presented at this stage, so that *hoc ferto* would in this context be out of place: the *fertum* has already been offered, and in so far as it has now been transferred to the possession of the deity being addressed the demonstrative *iste* is entirely appropriate. One may note that in *de agr.* 141, when it is a question of the first offerings proving unsatisfactory and substitutes being presented, the first offerings are designated with *ille* (*si quid tibi in illisce suouitaurilibus lactentibus neque satisfactum sit* 'if anything in that offering of *suouitaurilia lactentia* was not acceptable'), not with *iste*, precisely because they were not accepted by the god; the new ones being offered are, of course, referred to with *hic* (*te hisce suouitaurilibus piaculo* 'I (honour) you with this *suouitaurilia* as an expiatory offering').

There is then a perfectly good explanation for the occurrence of *iste* in the prayer in Cato that has nothing to do with rhyme or rhythm and everything to do with its function. It is not possible to explain **estu** in TI IIb23-4 in a similar way: the calf has not already been offered, and this is the first and only mention of it in direct address to the deity. Why then do we find, just on this one occasion, **estu** rather than a form from *esso*-? Are we to believe that on this one occasion the victim is considered from the point of view of the deity rather than the sacrificer? This is the interpretation of Ancillotti and Cerri (1996: 361): 'il punto di vista è nel destinatario, quindi nella divinità stessa ("codesto vitello" che tu divinità destinataria di questo mio messaggio hai davanti)'. This is hard to credit.

At this point we should look at the instances of *esto-* outside the Iguvine Tables to see if a similar interpretation appears plausible there. The reference to the gate in Po. 7/Um 5 (**estac vera**) appears in a building inscription, where all Oscan and Latin parallels would lead us to expect the *hic*-demonstrative (for Oscan, cf. **ekík sakaraklúm** (Ve. 150/Sa 7), **ekík pavmentúm** (Po. 133/Cm 4), **ekak víam** (Ve. 8/Po 1), **trííbúm ekak** (Ve. 11/Po 3), etc.), and Umbrian too has *eso* in the similar inscription on the cistern (Ve. 233/Um 8). There is simply no good reason, in relation to the gate, to accept here the claim of Rocca (1996: 55) that 'anche in questo caso, l'autore-esecutore considera l'oggetto iscritto rispetto all'utente scegliendo nel sistema il corrispondente di "codesto", ossia considerando la scritta pragmaticamente per la lettura da parte del destinatario ...'. Similarly with the gift inscription on the Pre-Samnite stamnos: all parallels suggest that *hic*-deixis is appropriate here; there is quite simply no occurrence of *iste* in such inscriptions in Latin epigraphy of the Republican period. Agostiniani (1979) does his best to make a case: 'lo stamnos immaginato, eccezionalmente, in prossimità del lettore non dell'autore del testo'; 'eccezionalmente' speaks for itself, and again this is not the most natural interpretation.

Agostiniani (1982: 26-7), in a more general discussion, speaks of the need to recognize that inscriptions involve both a Producer and a User, but it is a step too far to argue, as he does, that 'this' in an inscription must in consequence necessarily imply 'in prossimità dell'autore'. It seems rather that just as inscribed objects once spoke *in propria persona*, with first-person verbs and pronouns for self-reference, so they subsequently carry inscriptions designating themselves with demonstrative 'this', but with the same deixis, whence the regular choice of *hic* in Latin. 'Here lies X.', if found on a tombstone, does not imply proximity to the engraver but proximity to the funerary monument.

As for the South Picene examples, Marinetti (1985: 67-74), accepting the claims of Prosdocimi and Agostiniani, and noting that the local adverb **estuf** seems to be differentiated from **esmín** (on which see below, §11), argues again for second-person reference for the appropriate forms. In fact, since **estas** occurs in a broken and obscure context (see §7 above), it is only **estuf** 'here' that is usable; since the inscription in which it occurs seems to be funerary, the obvious parallels are with Latin epitaphs (*hic situs est* ..., etc.) and Paelignian *ecuf*, and parallels for the use of *iste* in such a context in Latin are completely lacking. There is one Latin *defixio* (*ILLRP* 1144) in which we find *mortuos qui istic sepultus est* 'the dead man who is buried there (near you)' and *ille mortuos quei istic sepultus est* 'that dead man who is buried there (near you)', where the local adverb *istic* is used, but the context here is quite different: rather than being an epitaph, this text is actually a prayer addressed to the god of the Underworld

(voc. *Dite Pater*), and the second-person reference is thus easily accounted for — the dead man is buried in the god's realm.

This is in fact the only instance in the Republican Latin inscriptions in Degrassi's collection of a form of *iste*. Even where an epitaph is addressed to a passer-by in the second person, reference to the monument itself is made with *hic*: e.g. *ILLRP* 808 *rogat ut resistas, hospes, t[e] hic tacitus lapis* 'this silent stone asks you, stranger, to stop'; consistent with this is the use of *eko-* in a like context in Paelignian, cf. Ve. 213/Pg 9 *eite uus ... puus ecic lexe* 'go, you who read/have read this'.

9. If one abandons the notion that *esto-* is functionally equivalent to *iste*, with second-person deixis, how then is the co-occurrence of the two demonstrative stems *es(s)o-* and *esto-* in Umbrian to be explained? How in particular are we to account for the fact that we find **estu vitlu** alongside *esu bue* in presentations of victims: what is the difference between the two passages that might explain the choice of demonstrative form? One obvious difference — obvious though not, I think, so far recognized as crucial — is that in the *tio(m) esu bue* formula we have an ablative, but in **tefe estu vitlu** an accusative.

If all the forms from *esto-* and *es(s)o-* are arranged by case, the following pattern emerges:

sg.				pl.		
	m.	f.	n.	m.	f.	n.
nom.		*eso*			**estas** (?)	
acc.	**estu**	**estam**	**este**			**estu**
gen.						**esumek**
abl.	*es(s)u*	*esa*				

with adverbial **estuf** 'here' and **esu(k)** 'as follows'.

Leaving aside for the moment nom. sg. f. *eso*, it can be seen that the remaining case-forms combine to give a paradigm showing an alternation between one stem (*esto-*) in the nominative and accusative and a second stem (*es(s)o-*) in the other cases. This immediately recalls the alternations that are found in the anaphoric pronoun **eio-/*eiso-* in both Oscan and Umbrian (cf. §1) and in the Oscan pronoun *eko-/ekso-* 'this' (cf. §2), and the similarity extends to the fact that the second stem shows a characteristic sigmatic form. Furthermore, the local adverb formed from the basic stem *eko-* in Paelignian *ecuf* is matched by the South Picene adverb **estuf(k)**, again formed from the same stem as the nominative and accusative.

The nom. sg. f. *eso* does, of course, present a problem, but it does not seem altogether implausible to postulate an analogical reformation on the basis of the oblique stem. An alternative possibility will be sketched below (§10) in connection with the derivation of the Umbrian stems.

Another possible difficulty might be adverbial **esu(k)**, if Meiser (1986: 119-20) is correct in taking this to be in origin the acc. pl. n. of *es(s)o-*, ending in *-ā*. He is certainly right in his contention that it cannot be an old abl. sg. m./n., since in the Latin alphabet it is consistently spelled <ESO>, while the abl. sg. m. singular of *es(s)o-* is consistently spelled <ESU>; and is likely to be right in his contention that a neuter singular **es(s)od* is to be excluded (*pace* Untermann 2000: 217) given that undoubted nom.-acc. sg. neuter forms of demonstratives in Sabellic all seem to show analogical extension of **-id* from the anaphoric pronoun (cf. Osc. **ekík**, U. **este**). But if a starting-point in a case-form with original **-ā* is sought, an alternative possibility would be that this is an old instrumental singular (feminine), such as is perhaps to be recognized in the Umbrian postposition **-ta**/*-to* 'from'.[5]

10. If one accepts that these forms make up a single paradigm, the question arises as to how the combination of stems is to be explained. Is the paradigm suppletive, and if so what is the source of the two stems and how did they come to be paired? Or are they somehow etymologically connected?

It seems quite possible that *esto-* is ultimately cognate with Lat. *iste*, even if there are difficulties in establishing a reconstruction: are we to postulate **esto-* or **isto-*? If the proto-form can be analysed at all, are we to recognize the pronoun **is* with an attached particle (with Umbrian *es-* under the influence of other pronouns with *e-*), or perhaps rather a particle **es* prefixed to the pronominal stem **to-* (with Lat. *is-* under the influence of *is*)?[6]

As for *es(s)o-*, it is almost universally regarded as identical to the Oscan *ekso-* stem, with assimilation of the consonant cluster (so, for instance, Buck 1928: 91, 142; Poultney 1959: 81, 110; Untermann 2000: 217-18). The attraction of this etymology, given the equivalence of function, is evident. The almost total restriction of *es(s)o-* to cases other than the nominative and accusative would also find a parallel in the Oscan paradigm. But it is not clear how the association with the quite different stem *esto-* would have come about, how the two stems

---

[5] See von Planta (1892-7: 2.454), Muller (1926: 222), and especially Brugmann (1897-1916: II/2.189, 702, 713, 787). Latin adverbs such as *intereā* may also continue old instrumental forms.

[6] See Leumann (1977: 470-1), Untermann (2000: 237). For an attempt to explain the origin of **es-* see Meiser (1998: 163).

that would presumably once have been distinct in function came to be combined and distributed within the paradigm according to the pattern attested in the texts.

Meiser (1998: 163) proposes to reconstruct parallel formations in Italic from *es-* with pronominal *so-* and *to-*, giving two stems, 'nebeneinander nur noch im Umbr. bezeugt, vgl. Abl. *ESSU* "hōc" bzw. **ESTU** "istum"'. One may note that the translations seem to presuppose a functional difference, which is here denied. It might nonetheless be tempting to pursue this line and start from a single paradigm that preserved the original distribution of *s-* and *t-* in the Indo-European pronoun, since this would provide an admirable explanation for nom. sg. f. *eso* (to IE *sā*) beside acc. sg. f. **estam** (to IE *tām*). For the oblique cases built on *es(s)o-*, however, something more is needed.

A possible solution would be to suppose that *esto-* and *es(s)o-* are related to each other in just the same way as Sabellic *eyo-* and *eiso-*, Oscan *eko-* and *ekso-*, in other words that *es(s)o-* represents the sigmatically extended form of *esto-*, viz. *estso-*, with assimilation.[7] The origin of the extension will be an enigma, just as it is for Oscan *ekso-*, but at least the paradigm, on this reconstruction, will share an overall pattern with other Sabellic pronouns.

11. A further question connected with demonstratives concerns U. **esmik** and *esme(i)*. The first of these forms is probably dat. sg. m./n., attested in the Iguvine Tables at Ia 28, 31, with what appears to be anaphoric function; the second may be dative or locative in VIa 5, 18 but is locative in VIb 55, and in all three of these instances the value of the pronoun seems rather to be 'this', as is reflected in most Latin versions of the Tables by the use of *hic*. In VIb 55, after an instruction to whoever belongs to certain foreign groups to depart from 'this *poplo-*' (*eetu ehesu poplu*), the text continues *nosue ier ehe esu poplu, sopir habe esme pople* ... 'if he does not go (*vel sim.*), if anyone catches him [or 'if anyone is caught'] in this *poplo-* ...', where the association with *esu* suggests similar deixis for *esme*. In VIa 5, 18 the phrase *esmei stahmei stahm(e)itei* comes in connection with observing birds immediately after the datives *mehe/tefe tote iiou(e)ine* 'for me/you, for the Iguvine state' and can either indicate a further beneficiary or designate the place of observation (the meaning of *stahmei stahm(e)itei* is unfortunately obscure); on any interpretation the pronoun can hardly be anaphoric, since this is not a reference back to something just mentioned (nor a reference forwards), and 'this' must be the appropriate translation. This all suggests that these forms have a basic meaning 'this', but can occasionally be used anaphorically (just like *hic* in Latin). This is consistent

---

[7] The precise form of the extension, and whether one should suppose syncope from something like *estVso-* (or from *ekVso-* for Oscan), does not affect the main point, which is the parallel with the Oscan paradigm.

with the evidence of South Picene: a possible dative **esmik** (RI 1), locatives (incorporating the postposition **-en*) **esmen** (CH 1, TE 2) or **esmín** (AP 1, MC 1, MC 2).[8] In these funerary texts, the forms most probably have the function 'this' (with reference to the monument) or 'here'.

For this reason, it seems best to follow Cowgill (1970: 140) in detaching these forms from the paradigm of the anaphoric pronoun (cf. Untermann 2000: 355-7), despite the seductions of a direct equation with Skt. *asmai*, Goth *imma*, and attaching them instead (while still recognizing pronominal endings **-smōi*, **-smei*) to the paradigm of *es(s)o-* 'this'. Cowgill postulated **ekso(s)m-* with syncope, but clearly a derivation from **estso(s)m-*, in line with the reconstruction of the paradigm suggested above, would work just as well.

12. If U. *esto-* and *es(s)o-* do constitute a single paradigm, and if they belong together in the way that has been suggested here, rather than originating as separate stems that were later combined, then there is no good evidence for a demonstrative pronoun with second-person deixis anywhere in Sabellic. This cannot, however, be taken as any sort of proof that Sabellic did not have a threefold system of deixis, just like Latin with *hic*, *iste*, and *ille*. The fact that Latin epigraphy of the Republican period, which is so much greater in extent than that of Sabellic, can offer only a single instance of a form of *iste* shows the danger of arguing from negative evidence in this connection.

References

Agostiniani, Luciano (1979). 'A proposito di Vetter 101 = *LIA* 20 D. Sui deittici in italico'. *SE*, 47: 386-8.
— (1982). *Le 'iscrizioni parlanti' dell'Italia antica.* Florence.
Ancilotti, Augusto, and Cerri, Romolo (1996). *Le tavole di Gubbio e la civiltà degli Umbri.* Perugia.
Brugmann, Karl (1897-1916). *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen* (2nd edn.). Straßburg.
Buck, Carl Darling (1928). *A Grammar of Oscan and Umbrian* (2nd edn.). Boston.
Conway, Robert S. (1897). *The Italic Dialects* (2 vols.). Cambridge.
Cowgill, Warren (1970). 'Italic and Celtic superlatives and the dialects of Indo-European', in George Cardona, Henry M. Hoenigswald, and Alfred Senn (eds.), *Indo-European and Indo-Europeans.* Philadelphia, 113-53.

---

[8] **]esmak toútaih** in R 1 is too uncertain in reading and interpretation to be useful here; for an explanation of **esmak** as a locative form, see Klingenschmitt (1992: 91).

Devoto, Giacomo (1962). *Tabulae Iguvinae* (3rd edn.). Rome.

Ernout, Alfred (1961). *Le dialecte ombrien*. Paris.

Klingenschmitt, Gert (1992). 'Die lateinische Nominalflexion', in Oswald Panagl and Thomas Krisch (eds.), *Latein und Indogermanisch*. Innsbruck, 89-135.

Leumann, Manu (1977). *Lateinische Laut- und Formenlehre*. Munich.

Marinetti, Anna (1985). *Le iscrizioni sudpicene*. Florence.

Meiser, Gerhard (1986). *Lautgeschichte der umbrischen Sprache*. Innsbruck.

— (1998). *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt.

Muller, Frederik (1926). *Altitalisches Wörterbuch*. Göttingen.

Planta, Robert von (1892-7). *Grammatik der oskisch-umbrischen Dialekte* (2 vols.). Strassburg.

Poccetti, Paolo (1979). *Nuovi documenti italici a complemento del Manuale di E. Vetter*. Pisa.

Poultney, James Wilson (1959). *The Bronze Tables of Iguvium*. Baltimore.

Prosdocimi, Aldo Luigi (1978). 'L'umbro', in A. L. Prosdocimi (ed.), *Lingue e dialetti dell'Italia antica*. Rome/Padua, 585-787.

— (1979). 'Le iscrizioni italiche. Acquisizioni, temi, problemi', in *Le iscrizioni pre-latine in Italia* (Atti dei Convegni Lincei, 39). Rome, 119-204.

Rocca, Giovanna (1996). *Iscrizioni umbre minori*. Florence.

Untermann, Jürgen (2000). *Wörterbuch des Oskisch-Umbrischen*. Heidelberg.

Vetter, Emil (1953). *Handbuch der italischen Dialekte*. Heidelberg.

The vocative's calling?

The syntax of address in Latin

Richard Ashdowne[*]

For centuries the vocative case has been the cause of amused puzzlement both to learners of Latin and learners of other languages. Winston Churchill famously recalled being introduced to the Latin language and coming across the vocative case for the first time:

(1)　　'What does it mean, sir?'

'It means what it says. Mensa, a table. Mensa is a noun of the First Declension. There are five declensions. You have learnt the singular of the First Declension.'

'But,' I repeated, 'what does it mean?'

'Mensa means a table,' he answered.

'Then why does mensa also mean O table,' I enquired, 'and what does O table mean?'

'Mensa, O table, is the vocative case,' he replied.

'But why O table?' I persisted in genuine curiosity.

'O table, – you would use that in addressing a table, in invoking a table.' And then seeing he was not carrying me with him, 'You would use it in speaking to a table.'

'But I never do,' I blurted out in honest amazement.

'If you are impertinent, you will be punished, and punished, let me tell you, very severely,' was his conclusive rejoinder.

(Churchill 1930: 25)

One can readily understand his astonishment, for it is indeed not often that one wants to address a table. Nevertheless, there are some important issues raised by looking at the phenomenon of the Latin vocative case, and it seems that traditional analyses of the vocative case do not provide a complete account of its use. This study is a preliminary investigation and is necessarily limited by the corpus on which it is based, namely Plaut. *Bacch.*, *Pseud.*, Ter. *Phorm.*, Cic.

Brut. and Petron. *Sat.* 28-78 (a total of 403 examples in approx. 60,000 words of text). However, the data cast doubt on the adequacy of some existing views, and several avenues are suggested for further investigation which could lead to a more helpful and precise analysis.

*

In order to study the use of the vocative,[1] first one must identify vocative forms. Classical Latin has a morphologically distinct vocative in only one class of nouns, namely second declension nouns in -us.[2] In all other singular nouns and all plural nouns the vocative is formally indistinguishable from the nominative. To isolate examples of these nouns in the vocative case (i.e., to be sure that a given example is vocative and not nominative), a number of different tests can be applied.

The case of a noun appearing in a phrase may be made clear by phrase-internal agreement shown on elements which do have a distinct vocative:[3]

(2)   *homo <u>lepidissume</u>* (Plaut. *Pseud.* 323)
       you dear delightful man

(3)   *<u>bone</u> vir* (Plaut. *Pseud.* 1145)
       good man

(4)   *<u>mi</u> pater <u>insperate</u>* (Plaut. *Rud.* 1175)
       dear father, father unhoped for

The interjection *o!* commonly accompanies addresses and exclamations; it is reasonable to assume that a form which could be either nominative or vocative is in fact vocative when accompanied by *o!*[4]

---

[1] By 'use' I mean the structures involved with and surrounding a vocative when we find one. I intend to say nothing about what determines whether one uses a vocative or not, nor do I intend to say anything about what one decides to call someone if one does use a vocative ('o great king!', 'my liege!' etc.). Both are fascinating sociolinguistic problems and have been discussed thoroughly by Dickey (1996). I limit my observations to the *grammatical* form of the relevant construction(s).

[2] In Plautus one also finds a vocative *puere* ('boy', e.g. at *Pseud.* 170, 241, 242, etc.), but second declension nouns in *-er* generally have vocative *-er*. The only other distinct vocative forms in Latin are in some nouns of Greek origin.

[3] Throughout this paper, I take examples from my corpus, supplemented with further material where necessary.

(5)   *o Troia, o patria, o Pergamum, o Priame ... senex* (Plaut. *Bacch.* 933)
       O Troy! O fatherland! O Pergamum! O old man, Priam!

Context can also make it possible to disambiguate a non-distinct case form. Where a doubtful case form appears close or adjacent to a second person pronoun in a clearly different case and cannot be explained in any other way, it is reasonable to analyse the form in question as vocative:

(6)   *vobis, mulieres, hanc habeo edictionem* (Plaut. *Pseud.* 172)
       Women, I have an announcement for you.

We can see, then, that it is possible to gather evidence for the use of the vocative case, not just where we find the *-e* or *-i* ending but sometimes also where the form is not distinct. Further morphological discussion beyond the principles for identifying examples is not essential for a consideration of the syntax (although observations about the syntax could eventually help to identify vocative examples). More detailed treatment of the issues together with consideration of the problem of indistinct second declension vocatives in *-us* (such as *meus*, *deus* and *populus*) can be found in Dickey (2000), Wackernagel (1912), Löfstedt (1956: 1.91-106), and Svennung (1958: 252).

What is worth pointing out, is that the internal syntax of a Latin phrase headed by a vocative element is unsurprising. A noun phrase in the vocative has an internal structure not discernibly different from that of a noun phrase in any other case.[5] For example, an attached genitive is perfectly possible:

(7)   *permities adulescentum* (Plaut. *Pseud.* 364)
       blight of the youth!

(8)   *senex minimi preti* (Plaut. *Bacch.* 444)
       you old man of no worth!

(9)   *homo nihili* (Plaut. *Bacch.* 1187)
       you worthless man!

Moreover, as we saw above, case agreement within the phrase (e.g. on qualifying adjectives etc.) is regular:

---

[4] A neuter noun in such circumstances could also be accusative and accordingly could be a *thing exclaimed over*.

[5] This said, there are semantic constraints which limit the evidence; vocative phrases are most commonly headed by a proper noun, which can rarely be the host for a genitive.

(10)   *bone serve* (Plaut. *Bacch.* 775)
       you good slave!

(11)   *mi pater insperate* (Plaut. *Rud.* 1175)
       dear father, father unhoped for

(12)   *Tiberine pater, te, sancte, precor* (Livy 2.10)
       O father Tiber, thee, holy one, I pray.

What continues to puzzle scholars, however, is what one might call the external syntax of the vocative phrase, i.e. how it is connected with and interacts with the syntax of any utterance with which it occurs. Before looking at Latin examples, it is instructive to consider the standard views of the phenomenon. (One word of caution is essential at this stage: many scholars use the terms 'address' and 'vocative' interchangeably and/or ambiguously to refer both to relevant case forms and to syntactic/pragmatic function. However, the distinction is absolutely critical. For the sake of precision, I use 'vocative' to refer to case forms and 'address' to refer to a pragmatic/syntactic function. In the following quotations, the distinction is not necessarily so rigorously observed.)

The most common view is that a vocative phrase is independent of any sentence which it may accompany:

(13)   Vocatives are ... an interesting grammatical category, again underexplored. Vocatives are noun phrases that refer to the addressee, but are not syntactically or semantically incorporated as the arguments of a predicate; they are rather set apart prosodically from the body of a sentence that may accompany them.

                                                                 (Levinson 1983: 71)

(14)   "Vocatives do not appear as dependents in constructions, but rather they stand outside constructions or are inserted parenthetically." (Hjelmsev 1935: 4) They are unlike other cases in that they do not mark the relation of dependents to heads. For these reasons, vocatives have not always been considered cases.

                                                                 (Blake 1994: 9)

It is worth dwelling on these views because they are, in a sense, no more than a *via negationis*. They exclude the vocative phenomenon from the explanatory domains of syntax and semantics, mainly on the theoretical grounds that the function of address does not parallel other functions which syntactic and/or semantic theory can encompass. Although this may seem intuitively reasonable,

it still does not in any way offer a good explanation for how the vocative phenomenon itself works.

Cautious advances towards such an explanation can be seen in the following:

(15) [T]he vocative [should] be recognized as a construction in which the dominant element is not case but person.

(Fink 1972: 65)

(16) [The] vocative marks a participant-role in the act of speech, whereas the other cases mark syntactic relationships between constituents of sentences.

(Vairel 1981: 440)

Such views point towards a potential analysis which is discourse-based, and we will return to some more detailed justification for an analysis on this level in due course.

First, however, let us compare actual Latin usage with a canonical description of the usage of the vocative:

(17) The Vocative stands apart from the construction of the sentence, with or without an Interjection.

(Kennedy [1888] 1946: §202)

In fact, a far more detailed taxonomy of vocative usage can be developed. The following is offered based solely on examples where distinctive vocative morphology is found; it applies equally well, however, to non-distinctive forms and might be useful in disambiguating them. The classification is not intended in itself to offer an analysis, merely to highlight the range of data which needs to be accounted for.

The first group I term 'isolative'. They are distinguished by the vocative phrase (NP) standing alone, without a sentence; it may, however, be strengthened with various particles (*o*, *euge*, *eheu*, etc.):

(18) *euge, homo lepidissume.* (Plaut. *Pseud.* 323)
      Wonderful, you dear delightful man!

A second group can be distinguished by the fact that they occur with a sentence which contains an explicit second person element; these I term 'quasi-appositional'. We can further subdivide this group into three types. There are examples where the vocative NP is adjacent to a second person pronoun or adjective:

(19) *nunc <u>Calidore te</u> mihi operam dare volo.* (Plaut. *Pseud.* 383)
      Now, Calidorus, I want you to assist me.

(20) *itur ad <u>te, Pseudole</u>.* (Plaut. *Pseud.* 453)
      They're coming at you, Pseudolus.

(21) *bene sit <u>tibi, Charine</u>.* (Plaut. *Pseud.* 714)
      Good luck to you, Charinus.

(23) *quo magis <u>tuum, Brute</u>, iudicium probo, ...* (Cic. *Brut.* 120)
      I therefore praise your judgement the more, Brutus, ...

(24) *id <u>tu Brute</u> iam intelleges, cum in Galliam veneris.* (Cic. *Brut.* 171)
      You, Brutus, will realise this presently, when you come to Gaul.

There are examples where there is no second person pronoun or adjective, but the vocative NP is adjacent to a second person verb:

(25) *quid <u>fles, cucule</u>?* (Plaut. *Pseud.* 96)
      Why do you weep, cuckoo?

(26) *<u>i, puere</u>, prae* (Plaut. *Pseud.* 170)
      Boy, go before me!

A final subgroup consists of examples where there is a second person element but the vocative NP is not immediately adjacent to that element:

(27) *<u>tuam</u> amicam video, <u>Calidore</u>.* (Plaut. *Pseud.* 35)
      I can see your girlfriend, Calidorus.

(28) *hoc <u>tibi</u> ille, <u>Brute</u>, minus fortasse placuit quam placuisset, ...* (Cic. *Brut.* 327)
      This perhaps pleased you, Brutus, less than it might have done, ...

(29) *nunc hoc <u>tibi</u> curandumst, <u>Chrysale</u>.* (Plaut. *Bacch.* 691)
      Now you'll have to sort this out, Chrysalus.

The third major group of examples is what I term 'directional'. These are where a vocative NP occurs with a sentence which contains no overt second person element:

(30)   *est misere scriptum, <u>Pseudole</u>.* (Plaut. *Pseud.* 74)
      It's woefully written, Pseudolus.

(31)   *odium igitur acerrimum patris in filium ex hoc, opinor, ostenditur, <u>Eruci</u>, quod hunc ruri esse patiebatur.* (Cic. *Rosc. Am.* 52)
      So then, I suppose, Erucius, this violent hatred of the father against the son is shown by allowing him to remain in the country.

The final group of examples consists of 'oaths'. These can themselves be isolative or with a sentence, and that sentence may or may not contain a second person element. However, what is important is that the vocative phrase is not coreferential with any second person element present in the sentence:

(32)   *<u>di immortales</u>, non Charinus mihi hic quidem, sed Copiast.* (Plaut. *Pseud.* 736)
      Ye immortal gods, I see this is not Chariness but Profusion!

(33)   *<u>o Zeu</u>, quam pauci estis homines commodi.* (Plaut. *Pseud.* 443)
      O Zeus, how few you men with the proper spirit are!

This last usage is not uncommon, but it is limited to a very small set of lexical items indeed. The two examples given seem to be the most common, and may be thought to have been lexicalized as interjections. One or two other proper names from myth were lexicalized as intensifying interjections (e.g. *edepol*, *ecastor*, *hercle*) probably via this stage.

The distribution of the examples in my corpus between these groups is: isolative 10.9%, quasi-appositional 66.5%, directive 20.6%, and oaths 2.0%.

Working from these four groups of vocative patterns, one can look at what connections, if any, exist between a vocative NP and a sentence with which it occurs. In turn, this will shed light on the kinds of structure which might explain the position of the vocative case in Latin syntax, and of the address function in a linguistic system. Inevitably the isolative configuration provides very little from which to proceed to a syntactic analysis, and the nature of the group of oaths (especially the possibility that they may be lexicalized) means that they too are not a good starting point.

Instead, we turn to the two other groups, whose patterns are shown by the figures to be frequent in the evidence. The questions that these raise are two. First, are the patterns (especially the high frequency of the quasi-appositional group) purely the result of chance? In other words, are these groups to be explained by analysing them as consisting of a sentence accompanied by some isolative vocative phrase? This is how they must be analysed on the basis of

Kennedy's description in (17) above, and indeed it is essentially how they would be analysed on the basis of the other standard *via negationis* views. Is it the case, then, that the use of the vocative case in general can be discussed without reference to a sentence with which it co-occurs?

The second question raised concerns whether the quasi-appositional group and the directional group can or should be analysed in the same way. A unified final account is highly desirable but the surface distinctions between these two configurations may be the result of significant structural differences and so may be informative about the structures involved.

<div align="center">*</div>

Let us now consider the matter of function. In drawing the terminological distinction between the vocative case and the address function, we raise a number of more fundamental questions concerning whether address is the only function that the vocative case is used for, whether address can be expressed in other ways, and most crucially what address actually is.

A way into tackling these questions is immediately apparent from the taxonomy I have just outlined. Vocatives, perhaps unsurprisingly, are often found next (or at least very close) to second person pronouns or adjectives; we can, therefore, investigate what other elements can occur in this kind of position or relation with respect to second person elements in order to establish either a parallel or contrast with the use of the vocative case.

In this position, we very often find the quasi-appositional vocative as I described above; however, we also find case agreement that is just like apposition in parallel first and third person constructions. In other words, there appears to be a syntactic alternation in these second person examples between strictly appositional case agreement on the one hand and quasi-appositional vocatives on the other.

(34)   *vides igitur ut ad <u>te oratorem</u>, Brute, pervenerimus ...* (Cic. *Brut.* 231)
       So, Brutus, you see that this has brought us all the way down to <u>you, the orator</u>, ...

(34') *sed <u>amico homini tibi</u> quod volo credere certumst.* (Plaut. *Bacch.* 1156)
       But <u>to a good friend like you</u> — I'm going to own up to what I want.

(35)   *nunc, <u>Calidore, te</u> mihi operam dare volo.* (Plaut. *Pseud.* 383)
       Now, I want assistance from <u>you, Calidorus</u>.

(35') *istactenu' <u>tibi, Lyde</u>, libertas datast | orationis.* (Plaut. *Bacch.* 168-9)
       Thus far <u>you, Lydus</u>, have been given freedom of speech.

(35") *principium, <u>Hedytium, te</u>cum ago ...* (Plaut. *Pseud.* 188)

    I'll begin with <u>you, Hedytium</u>, ...

Furthermore, the patterns are even more striking with adjectives; in general an adjective agrees in gender, number and case with the head it qualifies. (Strictly speaking, an agreeing adjective is not in apposition, but there are nonetheless justifiable parallels between the above examples of nouns in case agreement with the second person element and the examples below involving adjectives; the same applies *mutatis mutandis* to the examples above and below where case agreement is not present.)

(36)   *verum ego <u>te amantem</u>, ne pave, non deseram.* (Plaut. *Pseud.* 103)

        But never fear! I won't desert <u>you, my loving master</u>.

(36') *cavendumst mihi aps <u>te irato</u>.* (Plaut. *Pseud.* 473)

        I must beware of <u>you in your wrath</u>.

(36") *egone ut opem ferre putem posse <u>inopem te</u> mi?* (Plaut. *Bacch.* 638)

        Could I think that <u>you, unsupplied as you are yourself</u>, would be able to supply me?

(37)   *heus <u>chlamydate</u>, quid istic debetur <u>tibi</u>?* (Plaut. *Pseud.* 1139)

        Hey <u>you in the cloak</u>! What are <u>you</u> after there?

(37') *o <u>stulte, stulte</u>, <u>nescis</u> nunc venire <u>te</u>.* (Plaut. *Bacch.* 814)

        O you <u>poor poor idiot</u>, <u>you</u> don't know <u>you</u> are being sold this moment.

(37") *nisi scio probiorem hanc esse, quam <u>te, impuratissime</u>.* (Plaut. *Rud.* 751)

        But I do know that she's better than <u>you</u> are, <u>dirty beast</u>.

What is evident, then, is that there is variation in case usage. Moreover, this is not free variation but a consistent alternation between, on the one hand, elements that are semantically in apposition and, on the other hand, elements that are not. The above examples (34) and (36) show that the relation expressed by apposition corresponds to meanings such as '(in your state of) being, being on this occasion, in your capacity as'; it therefore affects the meaning of the sentence as a whole. The sentence would not mean the same if it were omitted. In contrast, the quasi-appositions seem to use any available description of 'you'; importantly, though, each adds nothing new to the meaning of the sentence it is accompanying because it does not alter the meaning of the second person element. The sentence would therefore mean the same if it were omitted.

    We can tell that this is the case because, for example, the apposition relation must refer to the whole set of individuals referred to by 'you' (which in the plural can include the addressee(s) *and* others) whereas the quasi-apposition applies only to the individual(s) so indicated; crucially this can lead to number

disagreement (as well as the case disagreement we have also seen) in the quasi-appositional construction, exemplified in (38) and (39) below — in apposition, however, case and number agreement are obligatory, as we saw in the examples above:

(38)   *ego <u>vobis, Geta,</u> alienus sum?* (Ter. *Phorm.* 545)
        But, Geta, am I not one of you (the family)?

(39)   *tum, igitur, dum licet dumque adsum, <u>loquimini</u> mecum, <u>Antipho,</u> | <u>contemplamini</u> me.* (Ter. *Phorm.* 549-50)
        So, then, while it is still possible and while I am still here, you (pl.) talk with me, Antipho, look at me.

As well as being semantic, apposition is uncontroversially also a syntactic relation and construction.[6] What I have dubbed quasi-apposition, however, is clearly altogether different semantically (its omission does not affect the meaning of the accompanying sentence); consequently we have no grounds here for treating it as syntactically the same either. This view is confirmed by the alternation in number agreement.

Indeed, if anything emerges from this alternation and distribution, it seems to be that the vocative in quasi-apposition stands outside the phrase with which superficially it appears to be connected. If we now bring into consideration the isolative and directional vocative groups, we can see that what unites all three groups is (part of) the traditional view, namely that the vocative case is indeed used to express address.[7] This function is quite evidently distinct from apposition, and, from the evidence we have seen, in Latin these functions have different surface realisations; they arguably therefore do not share an underlying structure.[8]

---

[6] I take it to involve the element in apposition being in the same phrase as the element it is in apposition to; more precisely, it is dominated by the maximal projection of that element, a state of affairs which has some significance for the semantics.

[7] There is no doubt that the isolative and directional vocatives express address inasmuch as they indicate the addressee; moreover, in quasi-apposition we can see that the vocative is referring to the addressee only, even when the second person element refers to others in addition.

[8] *Contra* Fink (1972: 65), who argues that the vocative is genuinely in apposition to a second person element, and that the vocative is the second person form of the noun which is invariable for case.

In fact, the independence of a vocative in quasi-apposition is such that it can appear outside a subordinate clause that contains the only second person element of the sentence:

(i)   '*Chrysalus mihi usque quaque loquitur nec recte, pater, |
        quia tibi aurum reddidi et quia non te defraudaverim.*' (Plaut. *Bacch.* 735-6)

Because we cannot tie the isolative and directional vocatives to any overt second person element (there being none present) and because furthermore we recognize quasi-appositional vocatives to be at the very least outside the phrase of the second person element they seem to accompany, it is easy now to see (a) why all three groups of vocatives can indeed be analysed together, if we wish, and (b) why they might be regarded as outside of the domain of the syntax of the sentence.

<div align="center">*</div>

We have seen some evidence to justify a view that the vocative is outside the normal explanatory domain of syntax (i.e. it isn't in apposition). However, in being apparently 'extrasyntactic' (or, better, 'extrasentential') we expect it to be subject only to a very limited number of constraints, none of which should be truly syntactic.

Other types of incontrovertibly extrasentential material include cries of pain (conventionalized linguistic ones like 'ouch') and parenthetical remarks. In some senses these seem to be altogether unconstrained. A cry of pain (whether non-linguistic 'aaarrghh!' or conventionalized 'ouch') can interrupt any unit of speech that is normally coherent, even words and very possibly syllables.[9] I am not aware of any phonological (or indeed pragmatic) constraint on the position, form or 'content' of such exclamations; however, in having reflex (i.e. subconscious) roots, this is hardly surprising. Nor am I aware of any internal grammatical structure within even conventional exclamations of pain.

A similar view might be taken of exclaimed urgent commands (e.g. 'watch out!', 'don't do that!') intended to warn someone of imminent danger. These, though made consciously (and having internal syntactic structure), may also interrupt even very small units, usually coming between words but conceivably interrupting them also.[10] They seem to have very limited positional constraints.

However, arguably neither of these categories need have (and in fact normally will not have) any connection with the utterance which it interrupts. This lack of connection between the interrupted utterance and the interjection or

---

Chrysalus keeps talking at me everywhere, and quite meanly, father,
all because I handed the gold over to you and did not defraud you.

[9] By 'interrupt' we must understand 'cause to be broken off before complete', i.e. it's not relevant whether the unit is resumed after the interruption. We are interested in the kind of constituents that are too fundamental to be broken in each case.

[10] It is worth bearing in mind that just because these kind of extra-sentential elements are *able* to interrupt smaller units, they need not necessarily do so and may often be found between words.

parenthetical is typical: there may coincidentally be a connection, but it is not required in order to license the presence of the parenthesis. The same is true of larger, more complex, parentheticals. These do tend to have an explanatory connection or relation, but it is not obligatory that they do:

(40)   The time has come — and I have seen the evidence — for all good men to come to the aid of the party.

(40')   The time has come — there are plenty of empty seats at the front — for all good men to come to the aid of the party.

However, leaving aside the cries of pain (which I take to lack internal structure), we can notice some patterns associated with parentheticals:

(a)   ability to interrupt freely;
(b)   absence of connection;[11]
(c)   no limit on the number present (beyond those concerned with interpretability).[12]

If we assume that the vocative is indeed syntactically and semantically free, as the traditional views claim and as our considerations earlier suggest, we might expect it to fall among these other extrasentential elements — and to obey the same very limited constraints. However, it is not clear that either vocatives or addresses are so unlimited.[13]

Specifically, addresses cannot interrupt freely, and do not have this complete freedom of placement. In Latin, for example, there are three typical positions for the vocative, namely sentence-initial, sentence-final and in second position in a sentence.[14] These make up 76% of the examples in the corpus. Of the remaining

---

[11] Under this heading we might also note that a parenthesis has to be self-contained. It cannot simply be divided into two or more parts and inserted discontinuously. One possible reason for this is the fact that it is not systematically connected with anything surrounding it, let alone with another parenthesis.

[12] Parentheses can occur between any pair of elements in a sentence and between as many such pairs as the speaker wants. However, a parenthesis between every such pair (or even between a large number of them) is ruled out by the same interpretability constraint that limits long sequences of genitive dependents, relative clauses, etc.

[13] In what follows I give English examples, conscious of the question begged. However, because I am arguing for constraints from the unacceptability of certain patterns, there are obviously no genuine Latin examples available. That said, I have found no Latin evidence that demonstrates the acceptability of any of the patterns that I claim to be unacceptable.

[14] By 'second position' I mean after the first coherent constituent (such as a noun phrase etc.), possibly with any immediately following unaccented words (unaccented *te* etc.).

Fraenkel (1965) uses the placement of the vocative to identify cola in Latin (and Greek) periods — his data and analysis strongly support the hypothesis that groups of words of a

24%, over half come between two clauses (usually the main clause and a subordinate clause, i.e. large syntactic constituents); smaller constituents such as noun phrases appear not to be broken by a vocative, although there are examples of parentheses breaking up even noun phrases:

(41) *...cum homine mirifico — ita mehercle sentio — Dionysio...* (Cic. *Att.* 4.11.2)
  ...with that wonderful man (for by Hercules that is how he seems to me) Dionysius...

What is most striking, however, is that the vocative when second usually seems to follow a topical constituent; this might suggest that in fact the vocative is in initial position before some topicalisation operation:

(42) *alii, Lyde, nunc sunt mores.* (Plaut. *Bacch.* 437)
  They're different, Lydus, the customs nowadays.

There remains considerable scope for investigating these patterns to see whether the vocative might be reduced to a single sentential position, but what is clear is that the vocative cannot be said to have the same degree of freedom of placement as cries of pain or even other (structured) parentheses.

Second, addresses actually do have some *necessary* connection with the accompanying utterance through something in the discourse context, viz. they must refer to the addressee(s) and are unacceptable if they do not. Since the addressee in question must specifically be that of the utterance they accompany, there is a consistent connection between the two *in a systematic way* — this cannot be said of the other parentheticals, where any connection is optional. We saw this most clearly in the quasi-appositional examples above: the referent of the second person element, while not limited to or determined by the vocative, cannot and must not exclude the addressee as named by the vocative. The acceptability of an address is determined by the utterance that appears to 'contain' it. We contrast, then, addresses with parentheticals whose acceptability is not dependent in this way on the surrounding utterance.[15]

Finally, the number of addresses is not unconstrained, in contrast to parentheticals. Let us consider (43) and (43'):

(43) The time has come, Mary, for all good men to come to the aid of the party.
(43') The time has come for all good men, my friend, to come to the aid of the party.

---

certain size and coherence (both syntactic and correspondingly phonological) cannot be broken up by an address.

[15] A parenthesis can even have a different addressee from that of the containing utterance, cf. (40').

'Mary' and 'my friend' are both fine alone as addresses, and there are at least these two positions in this sentence where an address is acceptable. The difference between address and parentheses is brought out in (44):

(44)   *The time has come, Mary, for all good men, my friend, to come to the aid of the party.

If the address is parenthetical in (43) and (43'), they should be able to be combined as in (44), which is unacceptable. If we assume for the sake of argument that addresses are parenthetical, we must treat 'Mary' and 'my friend' as two separate insertions, because parentheses may not be discontinuous (cf. n. 11 above).

Let us consider the two possible interpretations of (44) that I give in (45) and (45'):

(45)   *The time has come, Mary$_i$, for all good men, my friend$_j$, to come to the aid of the party.

(45') *The time has come, Mary$_i$, for all good men, my friend$_i$, to come to the aid of the party.

If (45) is said to only one person, we expect it to be ruled out as unacceptable because it is inconsistent with the discourse situation (a general pragmatic constraint). However, even when there are two suitable people present to be addressed, (45) is unacceptable. Furthermore, the coreferential alternative (45') ought (if addresses are parenthetical) to be fine, so long as at least one person is present to be addressed.

There appears, then, to be a constraint ruling out more than one address phrase.[16] We certainly do not want to claim that there can never be more than one parenthetical inserted in a sentence (so long as each could stand as a parenthetical alone in its own right), so what we have, then, seems to be some

---

[16] Contrast:

(ii)   The time has come, Mary$_i$, my friend$_i$, for all good men to come to the aid of the party.

Ignoring the punctuation, I take 'Mary, my friend' here to be a single address phrase insertion, rather than two separate insertions (which would have different intonation); the single phrase insertion is acceptable, but I suspect that it would be unacceptable with the double insertion intonation.

The acceptability of the following, then, depends on the acceptability of a conjoined address phrase without any (overt) conjunction:

(ii') $^?$The time has come, Mary$_i$, my friend$_j$, for all good men to come to the aid of the party.

kind of constraint that applies to addresses only and not to parentheses in general. We might reasonably conclude from this that addresses are simply not parentheses,[17] as we did from the 'connection' argument above.

It is hard, of course, to make the case for this uniqueness constraint in Latin, because Latin can allow phrases to be discontinuous. Even phrases which are undeniably phrases (and genuine syntactic constituents at that) can be split:

(46)   *non habes <u>venalem amicam</u> tu <u>meam Phoenicium</u>?* (Plaut. *Pseud.* 341)
       You don't have my girl Phoenicium up for sale, do you?

It is difficult to tell, therefore, whether the address in the following, for example, constitutes one or two phrases:

(47)   *o fortunate, cedo fortunatum manum, | Simo.* (Plaut. *Pseud.* 1065-6)
       O you lucky man, give me your blessed hand, Simo!

However, such examples are very few and far between; moreover, they never seem to occur in positions where both have to be taken with the same sentence (i.e. because they are both totally contained within it).[18] Usually, in fact, they can be resolved satisfactorily into two separate sentences by minor adjustments to the punctuation.

---

[17] In fact, this is a slight simplification. A sentence like (45) might be acceptable where the number of people present is more than one. However, the acceptability seems to derive from the conflation of (43) and (43') as a result of coordinated conjunction. Compare:
  (iii)   Here are your presents: I'm giving you this, Mark, you this, John, and you this, Mary.
  (iii') ... I'm giving you this, Mark, and I'm giving you this, John, and I'm giving you this, Mary.
I am less certain that (45') is ever acceptable, although certain ostensible addresses appear to be able to withstand such (repeated coreferential) insertion, e.g. 'Sir' or 'Ma'am'.

[18] We might note that this is one point where the group of oaths from the earlier taxonomy is relevant. It is possible to have both a vocative oath and a vocative address in Latin even in the same sentence; but if the oath is a lexicalized adverbial prayer rather than an actual address (i.e. this is fossilized vocative morphology), then this is unsurprising:
  (iv)   *pro di immortales, Chrysale, ubi mist filius?* (Plaut. *Bacch.* 244)
         Ye immortal gods, Chrysalus, where is my son?
In fact, it is worth pointing out that in the texts considered oaths occur exclusively in sentence initial position although we might have expected them to have a greater freedom of placement even than addresses. It is not clear to what extent their position reflects (a) the position in which they were typically found before they were lexicalized (and thus indicative of the unmarked or at least typical position for vocatives), (b) the position of such emphatic elements, and (c) other factors determining Latin word order.

A comparison, then, of addresses with parentheses and other obviously extrasyntactic/extrasentential elements has yielded some crucial differences. Taking the three broad differences between parentheticals and addresses together, we are driven towards a view which gives address some kind of structural relation to any sentence it accompanies. (I do not doubt, incidentally, that addresses can be parenthetical on occasions — but these differences suggest that it would be unwise to consider them always to be parenthetical and therefore constrained in this way purely through chance or coincidence.) What this relation appears to be, moreover, is a syntactic (and not just referential) one, there being a requirement for (a) coherence and (b) uniqueness of this element, both of which are useful diagnostic tests for syntactic constituents.[19]

*

A natural intuition, and one which we have used already, is to associate the use of the vocative case in some way with discourse. While all linguistic acts are in a sense inseparable from a discourse and its context, the address function is one which is solely concerned with articulation of these (e.g. initiating dialogue, marking turn-taking etc.). What is more, its link with morphology and consequently the use of forms according to syntactic rules is demonstrable. It is obvious, for example, that some pronouns or adjectives could never be used in address; furthermore, it seems unlikely to be a coincidence that these lexical items also lack a vocative form.[20] One instance of this is *tuus*, which necessarily refers to something other than the addressee inasmuch as it refers to something *belonging to* or otherwise *associated with* the addressee.[21] Another, perhaps more telling, example is *quis* and its adjective *qui* where the same morphological and semantic gaps are found.

---

[19] The very fact that languages like Latin have distinctive vocative morphology which is used in these systematic ways is a further argument in favour of a syntactic relation (just as the other morphological case forms in Latin express relations which are both syntactic and semantic).

[20] In Umberto Eco's *The Name of the Rose*, Brother William of Baskerville mocks mediaeval grammarians as the kind of people who would debate the significance of just such morphological niceties, but the question begged by such a debate is not an unimportant one:

(v) 'But those were times when, to forget an evil world, grammarians took pleasure in abstruse questions. I was told that in that period, for fifteen days and fifteen nights, the rhetoricians Gabundus and Terentius argued on the vocative of "ego" and in the end they attacked each other, with weapons.' (Eco [1980] 1984: 312)

[21] The absence of a vocative for *tuus* might conceivably be accidental (cf. English 'your honour', 'your majesty' etc.), but the importance of a discourse model cannot be overstated.

If we think about English, we note that one can ask a question about almost any element in a sentence. The constraints on which elements can be questioned, however, are germane to the vocative problem. Some adverbs, for example, can be questioned:[22]

(48)   She ran quickly down the street.
(48') How did she run down the street? Quickly.

However, so-called higher or sentential adverbs cannot be questioned:[23]

(49)   Luckily he ran down the street (which was why the bus didn't hit him).
(49') How did he run down the street? *Luckily.

In fact, those adverbials most closely tied to the discourse (e.g. speech act adverbials such as 'honestly, frankly' and evaluatives such as 'luckily, fortunately, happily' etc.) cannot even admit simple degree questioning:

(50)   Frankly they're going to win.
(50') How are they going to win? *Frankly.
(50") *How frankly are they going to win?

Many of these characteristics are shared by address phrases in English:

(51)   Mary, the door's open.
(51') *Who, the door's open?
(51") *Which Mary, the door's open?

One can question an address phrase but only indirectly (cf. matrix verb questioning for higher adverbials):[24]

---

[22] Cinque (1999) is a very full treatment of the syntax and semantics of adverbials, from which I summarize some relevant points.

[23] Both here and in the subsequent argument I naturally exclude the type of question which asks for the repetition of part of an utterance because the hearer suspects that it may have been misheard; such questions are always available to ask about any element whatsoever, even function words:
  (vi)   A: He's poor but he's honest.
         B: He's poor *what* he's honest?
         A: [I said, "He's poor] but [he's honest!]

[24] This test applies only to a limited degree to apposition, which is a relationship between two elements. Moreover, apposition is always optional and so a sentence is grammatical without it. To leave one of these elements in place and ask after some optional coreferential other

(52)   Who was/is told (that) the door was/is open? Mary.

(52')   How frankly were they told that they were going to win?

Cinque (1999) argues for a place within the syntax of a sentence for adverbials (even the sentential ones), and whether or not one accepts his particular structural analysis, it does not seem controversial to regard these elements as syntactically 'intrasentential' (integrated in and within the sentence). That address phrases parallel this pattern means that there is a convincing model from which to argue that they too (being discourse-associated elements) can be regarded as intrasentential, i.e. not completely independent of any sentence they accompany.[25]

In fact, to return to the contrast with parentheses introduced above, we can see that even the *possibility* of questioning a vocative in any way at all makes more explicit its connection with the accompanying utterance. A similar manner of question cannot be applied to parentheses, where a question involving the containing sentence is not systematically (if indeed ever) available — addresses (and so vocatives) on the other hand can only, it seems, be asked about using such a question.

*

To draw general conclusions about the issues involved in considering the vocative case and the address function is quite impossible at this stage. However, it seems that the evidence points towards a more refined analysis than before. Starting out from an elementary taxonomy of Latin data involving elements in the vocative case and the traditional 'theory' about these which places the vocative case and/or address function outside the realms of a sentence's syntactic structure, we have seen that it is possible to argue for some

---

element makes it unclear what kind of answer is expected; arguably the answer is already present in part in the question. Nonetheless, such a question becomes less bad if the field of apposition is indicated through a 'degree' question:

(vii)   You, the carpenter, open the door.

(vii')   *You, who, open the door?

(vii")   ?You, which carpenter, open the door?

[25] Of course, if an address were indeed a clause completely independent from the sentence, we would also expect it to be unable to be questioned. However, the similarity of patterning in matrix clause questioning suggests that the parallels should not be ignored. Further arguments in support of this parallelism concern the fact that speech act adverbials are cross-linguistically first in the ordering of higher adverbials, i.e. they must occur in the most peripheral adverbial positions in a sentence (typically as the first element, sometimes as the last); related to this is the fact that they do not appear (with the same meaning) in subordinate clauses. Address phrases have similar characteristics, some of which we have seen.

kind of structural integration in Latin which connects the use of the vocative case with the address function into a single position. A functional semantic alternation between address and apposition — in Latin made evident by morphological alternations between vocative and non-vocative elements —, then, can be paralleled and explained by what appears to be a syntactic alternation, between this address position and those positions usually accepted for appositional elements. The consequences of this theory remain to be spelled out in terms of formal syntactic structure, but the parallel with higher speech act adverbs may prove to be particularly relevant.

By combining the study of usage, meaning and function we have identified an outline for an overall account of the vocative phenomenon, at least for Latin, which comes out of a discourse-oriented approach. Moreover, we now see that the traditional or standard views were not in themselves wrong but that the situation is more complex that previously realized. No longer should the vocative case be a source of puzzlement, but a topic for renewed enquiry.

References

Churchill, Winston S. (1930). *My Early Life: A Roving Commission*. London.

Cinque, Guglielmo (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford.

Dickey, Eleanor (1996). *Greek Forms of Address: From Herodotus to Lucian*. Oxford.

— (2000). '*o egregie grammatice*: the vocative problems of Latin words ending in *-ius*'. *Classical Quarterly*, 50: 548-62.

Eco, Umberto ([1980] 1984). *The Name of the Rose* (tr. William Weaver). London.

Fink, Robert O. (1972). 'Person in nouns: is the vocative a case?'. *American Journal of Philology*, 93: 61-8.

Fraenkel, Eduard (1965). 'Noch einmal Kolon und Satz'. *Bayerische Akademie der Wissenschaften, Philosophisch-Historische Klasse: Sitzungsberichte*, 1965, 2. Munich.

Hjelmsev, L. (1935). *La catégorie des cas: étude de grammaire générale* (Acta Jutlandica, 7.1, I). Copenhagen.

Kennedy, Benjamin H. ([1888] 1946). *The Revised Latin Primer* (ed. J.F. Mountford). London, New York/Toronto.

Löfstedt, Einar (1956). *Syntactica* (2 vols., 2nd edn.). Lund.

Levinson, Stephen C. (1983). *Pragmatics*. Cambridge.

Svennung, Josef G. A. (1958). *Anredeformen. Vergleichende Forschungen zur indirekten Anrede in der dritten Person und zum Nominativ für den Vokativ*. Uppsala.

Vairel, Hélène (1981). 'The position of the vocative in the Latin case system'. *American Journal of Philology*, 102: 438-47.

Wackernagel, Jakob (1912). *Über einige antike Anredeformen*. Göttingen.

The type *faxo* in Plautus and Terence

Wolfgang David Cirilo de Melo[*]

1. Introduction

In classical Latin, most verb forms belong either to the *infectum*-stem, which denotes simultaneity, or to the *perfectum*-stem, which marks anteriority. In pre-classical Latin, on the other hand, there are a number of verb forms that do not fit into the regular paradigm based on these two stems: *attigas*[1] for example has a *perfectum*-stem, but the ending of the present subjunctive; and the stem of *duim* is neither that of the present nor that of the perfect.

Whereas most of these forms are already rare in early Latin and have become virtually extinct by Cicero's time, one group is still relatively frequent in the pre-classical writers; it consists of the types *faxo* (future), *faxim* (subjunctive), and *impetrassere* (future infinitive), the so-called 'sigmatic forms'.

They can be arranged in a paradigm as follows:

indicative:	*faxō, faxĭs, faxĭt, faxĭmus, faxĭtis, faxĭnt*
subjunctive:	*faxim, faxīs, faxit, faxīmus, faxītis, faxĭnt*[2]
infinitive:	*impetrāssere*[3]

In this paper, we shall be looking at the sigmatic indicatives, leaving out of account, however, *quaesso/quaessumus*[4] because they are semantically different, and forms like *moneris*, which are sigmatic only from a diachronic point of view.

As far as we can judge from our text corpus, the forms are used much less often around 160 BC than forty years earlier, and apart from a few exceptions like *ausim*, what we find in classical and post-classical Latin are deliberate archaisms rather than part of everyday speech. In Cicero's *De Legibus*, for

---

[1] Vowel length is marked only where necessary for the argument.
[2] *-int* is short because of Osthoff's law; the shortening of vowels before final consonant other than *s* is still in progress around 200 BC.
[3] The infinitive is restricted to a few verbs of the first conjugation; **faxere* is not attested.
[4] According to Leumann (1977: 623), *quaesso* has the same formant as *faxo*.

instance, they occur in legal phrases in imitation of old laws, while Apuleius is well known to draw on Plautine diction. For the present purposes, the comedies by Plautus (*c.* 254–184 BC) and Terence (*c.* 185–159 BC) have been chosen as a text corpus.[5] There are twenty more or less complete comedies by Plautus, while his *Vidularia* has been preserved only in part; in addition, there are a number of fragments from other plays. Terence's six comedies are all complete.

This choice of texts has several advantages: the comedies belong to the archaic period, a time when the sigmatic forms did not yet necessarily have a bookish air to them. Further, they give us at least a taste of diachronic development. And finally, the fact that we are dealing only with comedies reduces the risk that differences in usage between the authors are due to differences in genre or themes.

There is little doubt that the indicative forms mark some kind of future, but that is not sufficient. Three main questions arise:

(i)   Do they indicate tense or aspect? In other words, do they relate an event to a temporal reference point or do they mark it as being in progress, beginning, etc.?

(ii)   Do they belong to a specific register or can they be used in any circumstances?

(iii) How productive are the formations?

From a diachronic point of view, we shall want to know why the sigmatic forms die out in the way they do. A further question concerns their origin, but at present this will be left aside.

2. *The forms*

A look at the data reveals a surprising pattern of distribution: while many different verbs in all persons and numbers (except the third person plural, which may be chance) occur in subordinate clauses, the only form to be found in main clauses is *faxo*, the first person singular of *facere*.[6] The following two tables give the details:

	tokens in subordinate clauses	tokens in main clauses
Plautus	53	70
Terence	1	9

---

[5] We base our discussion on the editions by Lindsay for Plautus, and Kauer-Lindsay for Terence. Although there are some passages where I disagree with Lindsay's readings concerning sigmatic forms, these do not involve indicatives.

[6] The sigmatic subjunctives are not restricted in this way.

This first table displays the number of tokens in subordinate and main clauses. We can see that there are more tokens in main clauses than in subordinate clauses, especially in Terence. However, the first table has to be qualified by the second one:

	number of verbs in subordinate clauses	number of verbs in main clauses
Plautus	29	1
Terence	1	1

The second table shows that there are twenty-nine different verbs that have sigmatic indicative forms in subordinate clauses, but only one verb that appears in main clauses in such a form. And we must add that the verbs in subordinate clauses occur in all persons and numbers (with the exception of the third person plural), while the only form in main clauses is *faxo*.

This distribution requires an explanation. But we must first look at the semantics, register and productivity of the forms. We shall begin with the subordinate clauses and then examine the main clauses.

2.1. Subordinate clauses

The vast majority of sigmatic indicative forms in subordinate clauses occur in conditional clauses (forty-eight in Plautus, one in Terence); only a few forms are found in relative clauses (two in Plautus) and in temporal clauses (three in Plautus).

As we have seen, it is generally acknowledged that these forms have future reference, but that in itself does not tell us much about their semantics. However, the sigmatic futures in subordinate clauses are often co-ordinated with future perfects, as in (1):

(1) Plaut. *Aul.* 56-9 (Euclio is threatening his servant):

EUCLIO: *Si hercle tu ex istoc loco*
  digitum transvorsum aut unguem latum EXCESSERIS
  aut si RESPEXIS, *donicum ego te iussero,*
  continuo hercle ego te dedam discipulam cruci.

EUCLIO: By Hercules, if you WILL HAVE GONE[7] a finger's or a nail's breadth from your place, or if you WILL HAVE LOOKED BACK until I've told you, I will, by Hercules, immediately put you on the cross for a lesson.

---

[7] The translations seek to render the Latin tenses as literally as possible, even though this often results in unidiomatic English.

In (1), the future perfect *excesseris* is co-ordinated with the sigmatic future *respexis*. It is unlikely that *respexis* has a temporal reference different from that of *excesseris*, since it merely introduces a more specific condition, not a new one. Similar co-ordination of sigmatic future and future perfect in conditional clauses can also be found in Plaut. *Bacch.* 847-9, Plaut. *Cas.* 1001-3, and Plaut. *Rud.* 775-6. Compare also the clauses in (2):

> (2) Plaut. *Fretum* (the speaker is in a catch-22 situation):
> PERIBO si non FECERO: si FAXO, VAPULABO.
> I WILL PERISH if I WILL not HAVE DONE it; if I WILL HAVE DONE it, I WILL BE
> BEATEN UP.

Here, the two conditional and main clauses are not co-ordinated, but parallel to each other and have the same time reference; the future perfect or sigmatic future stands in the conditional clauses, and the simple future in the main clauses. Other instances of parallelisms in conditional or relative clauses can also be seen in Plaut. *Capt.* 122+124, Plaut. *Cas.* 1015-18, Plaut. *Mil.* 156-7+163, and Plaut. *Pseud.* 531-4.

These instances of co-ordination and parallelism seem to indicate that the sigmatic futures have the same temporal reference as the future perfects in subordinate clauses. The future perfect is used to show that the event in the subordinate clause is over before the one in the main clause begins. Consequently, the sigmatic futures are also likely to indicate anteriority. Just like the future perfects, they can all be interpreted as perfective,[8] but unlike them, they are restricted to telic verbs.[9]

While it is true that we expect the protasis of most conditional clauses to refer to events anterior to those mentioned in the apodosis (and similar considerations may be made for other subordinate clauses), the anterior reference of our sigmatic forms does not necessarily depend on the subordinate clauses to which they belong. In conditional clauses we also find simple futures which do not express anteriority:

> (3) Plaut. *Epid.* 263-4 (Epidicus is proposing a plan):
> EPIDICUS: *Immo si PLACEBIT, utitor,*
>   *consilium si non PLACEBIT, reperitote rectius.*

---

[8] Under perfectivity I understand that an event is described as a whole, not as being in progress and incomplete.

[9] In Plaut. *Cas.* 1001 *amare* seems to mean 'make love' rather than stative, atelic 'love'. Sigmatic subjunctives need not be telic: compare *licessit* in Plaut. *Asin.* 603.

EPIDICUS: Well, if you WILL LIKE it, use it; if you WILL not LIKE the plan, find a more
             suitable one.

In (3), the addressees must still like the plan when the main clause action begins.

In other words, the co-ordination of sigmatic forms and future perfects is significant. We may see whether an anterior meaning is suggested or contradicted by other evidence. Unfortunately, the temporal adverbials are not very helpful. *Posthac* (Plaut. *Cas.* 1001) merely situates the action in the future, and *praeterhac* (Plaut. *Men.* 113, Plaut. *Rud.* 1118, Plaut. *Stich.* 345) need not even be taken as temporal — it can have a discourse function as well. *Ante* (Plaut. *Asin.* 818) and *priu'* (*... quam*, Plaut. *Epid.* 122-3[10]) suggest anteriority, even if they themselves are not sufficient to prove it. Yet none of the occurrences of the sigmatic indicatives points in other directions, and we must conclude that these forms mark future anteriority.

If this is so, how do the sigmatic futures differ in usage from the future perfects? Conceivably the difference is not one of meaning, but of register. It has been shown by Haffter (1934) that, while Terence uses a relatively uniform language, the language of the Plautine *cantica*[11] contains more elements of elevated language than that of the *senarii*. According to Happ (1967: 88), the ratio of *cantica* and *senarii* is 3:1. We shall only look at Plautus because there is just one token in Terence. The two tokens in relative clauses are in *cantica*, and of the three tokens in temporal clauses, the one that is not in a *canticum* is in a parody of a law (*Persa* 70), where a higher register is appropriate.

At first sight, the situation seems to be different in conditional clauses: of the forty-eight tokens in Plautus, thirty-four occur in *cantica* and fourteen in *senarii*, which is the distribution that we would expect for stylistically unmarked forms. However, when we look at these fourteen forms more closely, we can see that most of them are found in contexts where Plautus normally uses elevated language:

(i)   invocations of gods (*Bacch.* 848, *Rud.* 1348); compare (4):

(4) Plaut. *Bacch.* 847-9 (a soldier is pretending to be angry):
CLEOMACHUS: *Nam neque Bellona mi umquam neque Mars creduat,*
             ni illum exanimalem FAXO, *si convenero,*
             nive exheredem fecero vitae suae.

---

[10] *Hodie* in the same passage may be temporal or may express anger: compare Donatus on *Ad.* 660.

[11] Under *cantica* I understand anything that is not in *senarii*, in other words, both sung and recited passages.

CLEOMACHUS: For never shall Bellona or Mars trust me any more, if I SHALL not HAVE MADE him dead, if I've found him, or if I shall not have disinherited him of his life.

(ii)   legal language (*Asin*. 770, *Asin*. 794, *Rud*. 811[12]);
(iii) asides by someone overhearing a conversation (Happ 1967: 182) and moralizing soliloquies (*Asin*. 818, *Stich*. 192);
(iv)   passages spoken by 'tragic' or boastful characters (*Capt*. 124: Hegio, *Capt*. 695: Tyndarus, *Epid*. 441: Periphanes, a former *miles gloriosus*).

Four tokens remain. Among these, the one in *Poen*. 428 is clearly in a passage belonging to a higher register: the speaker, Agorastocles, literally heaps up figures of speech (*geminatio*, parallelism, and a failed attempt at an *adynaton*). Similarly, the form in *Persa* 393 is prefaced by the bombastic statement *pol deum virtute dicam et maiorum meum*, 'by Pollux, I will speak by the excellence of the gods and of my ancestors' (*Persa* 390). *Pseud*. 533 is preceded by a hyperbolic comparison between the slave Pseudolus and king Agorastocles. And finally, *Aul*. 58 is in a passage where Euclio, one of the main characters, is introduced. In such contexts a higher register is not unusual.

Although we cannot be absolutely certain about all the passages listed above, most of them appear to belong to a higher style. If so, we can reasonably conclude that the sigmatic indicative forms in subordinate clauses belong to a higher register.

This may be interpreted in different ways. At one extreme it may signal complete fossilization. Is this a likely interpretation for the forms in Plautus and Terence? In relative and temporal clauses they are indeed so rare that they can hardly be called productive in Plautus, and in Terence they do not occur at all in these contexts. However, in conditional clauses there are forty-eight tokens in Plautus (as opposed to the one form in Terence). This is a considerable number; can we therefore call the sigmatic indicatives productive in Plautine conditional clauses? They do not occur in formulaic phrases, and the forty-eight tokens belong to twenty-seven different verbs. (Seven tokens belong to *facere* and two to a compound of it, *ecficere*.) This distribution does not speak for fossilization. We may qualify this assertion by comparing the figures for the sigmatic indicatives to those for the future perfect, which, as we have seen, has the same meaning, but belongs to a different register:[13]

---

[12] Compare Marx (1959: 160, *ad loc*.).

[13] Many forms are ambiguous between future perfect and perfect subjunctive. I have therefore only counted a subset of instances in which there must be a future perfect, namely if the conditional clause depends on a clause with a finite future form, on a perfect with future

	si[14] + sigmatic fut.	*si* + fut. perf.	*nisi* + sigmatic fut.	*nisi* + fut. perf.
Plautus	36	123	12	27
Terence	1	32	0	1

We can see that the future perfect, which is of course not restricted to conditional clauses, is far more productive. But it is likely that the sigmatic forms had a limited productivity in conditional clauses with a higher register in Plautus. In Terence, on the other hand, the formation is fossilized.

2.2. Main clauses

The only form that occurs in main clauses is *faxo*: there are forty-one instances in Plautine *cantica*, twenty-nine in his *senarii*, and nine in Terence. It is found in different construction types, for example in the type *faxo hic aderit*, which could provisionally be translated as 'I shall bring it about that he shall be here'. In other words, this is a causative construction.

Typical examples are (5) and (6):

(5) Plaut. *Epid*. 656 (a slave wants his master to be patient; he will reveal everything a bit later):

SERVUS: *Cetera haec posterius* FAXO SCIBIS *ubi erit otium.*

SLAVE: I'LL TAKE CARE that you WILL KNOW the other things later when we are free.

(6) Plaut. *Truc*. 643-4 (a soldier is angry because a courtesan has not treated him well; he is threatening to take revenge):

MILES: *Ego* FAXO DICAT *me in diebus pauculis*
            crudum virum esse.

SOLDIER: I WILL SEE TO IT that she SAYS in a few days that I am a savage man.

The verb with which *faxo* is construed can be (i) a simple future as in (5) (28 times in *cantica* and 20 times in *senarii* in Plautus, 5 times in Terence), (ii) a present subjunctive as in (6) (9 times in *cantica* and 5 times in *senarii* in Plautus, twice in Terence), (iii) a future perfect (3 times in *cantica* and twice in *senarii* in

---

reference like *perii*, on a gerund construction, or on an imperative or a directive expression/a wish like *divi me perduint* 'may the gods destroy me'. I have, on the other hand, counted all the instances of sigmatic futures, not just a subset. The results concerning the ratio of sigmatic future and future perfect are still obvious.

[14] I include *sive* and other alternatives for *si*, just as I include *nive* and alternative expressions under *nisi*. The following future perfects or perfect subjunctives have not been considered: *compegerint* (Plaut. *Amph*. 155), *praehibuerit* (Plaut. *Merc*. 1023), *feceris* (Plaut. *Trin*. 348), *habuerit* (Ter. *Hec*. 718), *redierit* (Ter. *Phorm*. 248), *eiecerit* (Ter. *Phorm*. 627).

Plautus), and (iv) a perfect subjunctive (once in a *senarius* in Plautus). In addition, there is one instance of *faxo ut* (Plaut. *Asin*. 897), and there is also *faxo* with a double accusative (once in a Plautine *senarius* and twice in Terence).

If we compare these constructions with those of classical Latin, the obvious contrast is in the use of *ut*: *facio ut* + subjunctive is the normal causative formation of Latin, while *ut* is missing here. In pre-classical Latin, however, we have many sequences of verbal forms followed by subjunctives without *ut*. Particularly frequent is the type *fac* + subjunctive, where the imperative *fac* alternates with *fac ut*; less frequent, but attested is the future *faciam* + subjunctive,[15] which alternates with *faciam ut* + subjunctive. In other words, in the constructions of *faxo* we must assume that the forms without *ut* are the original ones, while the one example with *ut* is an innovation based on *faciam ut*. The striking phenomenon is not the absence of *ut*, but the presence of a future or future perfect in the slot otherwise occupied by the subjunctive. This could perhaps be explained if *faxo* began as a parenthetic causative construction.

Are there any semantic differences between the constructions? As far as we can see, *faxo* + simple future and *faxo* + present subjunctive are not distinct from each other. They occur in the same contexts, and there are no restrictions as to what verbs can be governed by *faxo*. We find both *faxo scibis* and *faxo scias* ('I shall make you know'). In other words, *faxo* can be combined with stative verbs. However, we also encounter both *faxo dices* and *faxo dicat* ('I shall make you/him say'). Thus *faxo* can also be combined with telic events.

Just as *faxo scias* means 'I shall bring it about that you know', the only instance of *faxo* + perfect subjunctive (Plaut. *Men*. 521) is of the type 'I shall bring it about that you have done'. Of course a future action such as *faxo* cannot cause a past event; but it can be the cause of results of the past action, and in our example this result is punishment, expressed in an obligatory adjectival phrase *haud inultus* ('not unpunished').

In those instances where *faxo* governs a future perfect, this future perfect has the value it would also have in independent main clauses. Thus, the future perfect indicates the result of the action expressed by the verb in (7):

(7) Plaut. *Pseud*. 766 (Pseudolus is trying to fool a pimp and is comparing this to capturing a city):

PSEUDOLUS: *Iam ego hoc ipsum oppidum* EXPUGNATUM FAXO ERIT *lenonium.*

PSEUDOLUS: I'LL SEE TO IT that this pimp-town WILL BE TAKEN soon.

---

[15] The two instances of a first person future in our corpus are Plaut. *Amph*. 63 (*faciam sit*) and Plaut. *Amph*. 876 (*faciam … fiat*).

Here, Pseudolus uses the future perfect not to indicate that he will do something, but rather that he will have done it and that the results will then be obvious to everyone.

In the other four instances, the future perfect is employed to assess the consequences of an event and does not introduce a new action.[16] Compare (8):

(8) Plaut. *Trin*. 59-60 (Megaronides thinks his wife is even worse than Callicles'):
MEGARONIDES: *Vin commutemus, tuam ego ducam et tu meam?*
> *FAXO hau tantillum DEDERIS verborum mihi.*
MEGARONIDES: Do you want us to swap, that I take your wife and you mine? I WILL BRING
> IT ABOUT that (thereby) you SHALL not HAVE FOOLED me at all.

In (8), the future perfect *dederis* does not introduce any new action. The clause is an assessment of the consequences of the event described in the previous clause.

Note that *faxo* is never combined with a first person singular or plural; thus, the types **faxo mactabo* and **faxo mactemus* are absent.[17] There seems to be a constraint against the causer being identical to the causee, but this is not general, as we do have *fac adsis* etc. However, a causative interpretation is not always possible; compare (9):

(9) Plaut. *Men*. 791 (a father is scolding his daughter because she wants to control her husband, who loves a courtesan):
SENEX: *Ob istanc industriam etiam FAXO AMABIT amplius.*
THE OLD FATHER: He WILL CERTAINLY MAKE LOVE to her all the more because of your
> watchfulness.

Here the causative translation 'I shall bring it about that' would be absurd. A father cannot want his daughter's husband to love a courtesan. The father is not the causer — he has nothing to do with his son-in-law's love for another woman. In fact, the cause is his daughter's behaviour. In contexts like this, *faxo* has been reanalysed and can be translated adverbially as 'certainly'.[18] It has

---

[16] For this value of the future perfect compare Risselada (2000) and also Gaffiot (1933). In the English translation, we can always add 'thereby' to makeit clear that no new event is introduced.

[17] Similarly, *faciam ut* + first person is rare as a causative construction. Plaut. *Capt*. 385-7 (*faciam sedulo, ut … id petam*) is not really an exception because the *ut*-clause is an explanatory adverbial clause, so that *faciam ut* cannot be taken causatively. Ter. *Hec*. 244-5 could be given a causative interpretation.

[18] It seems that the type *faxo erit* expresses certainty, while the type *faxim sit* (cf. Plaut. *Trin*. 221) marks possibility.

become a marker of certainty. There are two ways in which we can interpret this: the first is that in contexts like these, *faxo* has been re-analysed syntactically and has become an adverb (in *faxo* + future) or a 'subordinator' governing the subjunctive (in *faxo* + subjunctive). The word 'subordinator' is not a very good one, though, because subordinators introduce subordinate clauses that cannot stand on their own, while *faxo* does not do so. But we can compare *faxo* to *forsitan* ('perhaps').[19] Alternatively, *faxo* could have been re-analysed purely semantically and could have become a parenthetical expression meaning 'I assume'.[20]

In many cases, *faxo* can be interpreted as either causative or non-causative: if the speaker wants to be seen as the causer, *faxo* is verbal; otherwise it is a marker of certainty. But in that respect most examples are ambiguous because the speaker has a choice. If he says *iam faxo hic aderit*, he may want to emphasize his involvement as the causer, in which case *faxo* would be causative; or he may wish to focus on the event itself and the certainty of its coming about rather than on his own activities, and then *faxo* would be a marker of certainty, be it verbal or not.[21] In some cases one interpretation is more likely than the other, for instance in (10):

(10) Plaut. *Curc*. 586-7 (a soldier is looking for a man nicknamed Curculio — 'weevil'):

Miles: *Ubi nunc Curculionem inveniam?*

Cappadox: *In tritico facillume,*
             vel quingentos curculiones pro uno FAXO REPERIAS.

---

[19] Krebs and Schmalz (1905: 603) describe *forsitan* as a particle governing the subjunctive. In Cicero, it is always combined with a subjunctive. The subjunctive often seems to be motivated solely by the presence of *forsitan*; after all, *fortasse* has a similar meaning, but is also combined with the indicative. However, as there is an instance of *forsitan* + indicative as early as Lucretius (5.104), it could in theory be argued that the subjunctives in Cicero are independent of *forsitan*. This line of argument is impossible when a synchronically opaque particle governs a construction such as the *accusativus cum infinitivo*, which *fortasse* sometimes does in Plautus. Compare *Poen*. 1004: *fortasse medicos nos arbitrarier* 'perhaps he thinks we are doctors'.

[20] Both analyses have advantages. *Faxim* is verbal — it can be combined with *ego* — and means 'I assume', so the semantic re-analysis of *faxo* is not problematic in that respect. On the other hand, such parenthetic modalizing expressions in Latin are normally in the present tense, not in the future. Non-parenthetic modalizing verbs normally govern the *accusativus cum infinitivo*.

[21] A similar situation may obtain in ordinary clauses with a transitive verb: the speaker sometimes prefers not to state that he is the agent and therefore uses the passive. Cf. Plaut. *Mil*. 24: *epityra estur insanum bene* '*epityra* can be eaten terribly well'.

SOLDIER: Where am I now to find Curculio?

CAPPADOX: Most easily in the wheat. You WILL
                CERTAINLY FIND even five-hundred weevils instead of one.

In (10), Cappadox can hardly be seen as the causer, so we interpret this as a marker of certainty. But (9) and (10) are the only examples where the semantics of *faxo* is clear in this respect.

Sometimes the context can give us hints on how to interpret *faxo*, for example in (11):

(11) Plaut. *Men.* 950-1 (a doctor is telling Menaechmus how he wants to cure him):

MEDICUS: *Elleborum POTABIS FAXO aliquos viginti dies.*

MENAECHMUS: *AT EGO te pendentem FODIAM stimulis triginta dies.*

DOCTOR: I WILL SEE TO IT that you WILL DRINK hellebore for some twenty days.

MENAECHMUS: BUT I WILL PROD you, while hanging, with goads for thirty days.

In (11), *faxo* ought to be taken as causative because Menaechmus replies *at ego* and thus contrasts himself with the doctor; this can only make sense if the doctor uses *faxo* not as a marker of certainty, but as a causative verb with himself as the subject. For similar reasons, *faxo* is probably causative in Plaut. *Asin.* 132, Plaut. *Poen.* 1228, Plaut. *Truc.* 761, and Ter. *Ad.* 847; the speakers are emphasizing that they will *cause* something.

If *faxo* as a marker of certainty is not verbal any more, we can look at syntactic criteria as to where *faxo* has to be verbal and thus causative: in *faxo ut*, *faxo* is verbal because otherwise there could be no subordinator *ut*. And when there is *ego faxo* instead of mere *faxo*, this must be verbal — otherwise it would make no sense to express a subject *ego*. *Ego* is used eighteen times. Similarly, when *faxo* is combined with a double accusative it has to be verbal, or the clause would have no verb at all. But if *faxo* as a marker of certainty is verbal, only *faxo* + double accusative has to be causative from a syntactic point of view.

What is the temporal meaning of *faxo*? Again, temporal adverbials do not tell us much. The following adverbials occur in Plautus and Terence[22] (their frequency is given in parentheses): *actutum* (1), *cras* (1), *extemplo* (1), *hau multo post* (2), *hodie* (3[23]), *iam* (25), *in diebus pauculis* (1), *nunc* (1), *post* (1), *posterius* (1), *posthac* (1), *priu' ... quam* (1), *quom audiet* (1), *temperi* (1), *tum* (1), *ubi erit otium* (1).

---

[22] We exclude the adverbials of frequency (*saepe*, Ter. *Eun.* 285) and duration (*aliquos viginti dies*, Plaut. *Men.* 950). In any case, they seem to modify the verb combined with *faxo* rather than *faxo* itself.

[23] Except for Plaut. *Mil.* 1367, these may again express anger rather than time.

The problem with most of the adverbials is that in a phrase like *iam faxo hic aderit* (Ter. *Phorm*. 308) it is not clear whether *iam* modifies *faxo* or *aderit*. There are only two examples where we can be sure that the adverbial belongs to *faxo*: Ter. *Haut*. 341 and Plaut. *Rud*. 800. In Ter. *Haut*. 341, *faxo* governs a double accusative rather than a finite verb, and consequently *iam* must modify *faxo*. The same is true for Plaut. *Rud*. 800, where the double accusative forces us to take *hodie* with *faxo* — yet *hodie* could express anger instead of time: compare Donatus on *Ad*. 660. In short, the temporal adverbials do not help us much.

We used 'I shall bring it about that' as a provisional translation for *faxo*. In subordinate clauses, the type *faxo* can always be interpreted as anterior; in main clauses, no such interpretation is possible, cf. (6) above. In this example, the soldier simply refers to the future; it is impossible to detect any anteriority. The example can be translated perfectively, but not imperfectively: 'I'll be in the process of making her say' does not make any sense. However, this perfective interpretation can simply be the result of the kind of speech act we are dealing with: *faxo* is used in promises and threats, and we normally promise or threaten to do things completely, from beginning to end.

We learn more from the tenses with which *faxo* is co-ordinated.[24] In (12) it is the simple future:

(12) Plaut. *Asin*. 131-2 (Argyrippus is angry with a procuress):
ARGYRIPPUS:*Ibo ego ad trisviros vostraque ibi nomina*
FAXO ERUNT, capiti' te PERDAM ego et filiam.
ARGYRIPPUS: I WILL GO to the *tresviri* and TAKE CARE that your names WILL BE there, I WILL DESTROY you and your daughter completely.

Here, *faxo* is co-ordinated with two simple futures, *ibo* and *perdam*. *Faxo* in Ter. *Ad*. 847 is co-ordinated with a future perfect *videro*, but *videro* is a fossilized idiom that does not indicate anteriority; besides, *faxo* is parallel to the following *faciam*. In (13) it is not clear if there is any co-ordination between *faxo* and *faciet*:

(13) Plaut. *Poen*. 371-2 (Milphio has to pacify Adelphasium):
MILPHIO: *Ego FAXO, si non irata's, ninnium pro te DABIT*
ac te FACIET ut sies civis Attica atque libera.

---

[24] Only those cases in which *faxo* has to be verbal were considered.

MILPHIO: I WILL TAKE CARE that, if you are not angry, he WILL PAY for you, little horse, and he WILL TAKE CARE (or: and THAT he WILL TAKE CARE) that you are a free Athenian citizen.'[25]

In (13), *faxo* and *faciet* could be regarded as being co-ordinated. Alternatively, it could be *dabit* and *faciet* that are co-ordinated. The sentence is ambiguous. In the latter case, there would be a double causative construction: *faxo faciet ut*, with two causers, the speaker and his master, who is the subject of *faciet*. In any case *faxo* is clearly a future. The general conclusion is that in main clauses *faxo* has simple future meaning and does not correspond to a future perfect. This is a crucial difference to the sigmatic forms of the subordinate clauses.

What about the register of *faxo* in main clauses as compared to that of the forms in subordinate clauses? We can immediately see that *faxo* is disproportionately frequent in Plautine spoken verse, which indicates that it is not marked for a particularly high register. This conclusion becomes even more certain when we consider that all types of character use it indiscriminately. The same can be said about Terence's usage, where *faxo* is found in ordinary as well as stylistically marked passages.

With regard to the productivity of sigmatic forms in main clauses, we have already said that the only form to be found is *faxo*. There are some recurrent formulae in Plautus: among the forty-eight tokens of *faxo* + simple future, a form of *esse* occurs eleven times, a form of *scire* nine times, and a form of *adesse* four times. Similarly, *dicere* is found three times in the future and three times in the subjunctive. However, twenty-four verbs are found in the simple future and eleven in the present subjunctive; so we are not dealing with fixed expressions. The same is true of Terence; there are far fewer tokens, but the only form to occur twice is *aderit*.

There are seventy tokens of *faxo* in main clauses in Plautus, but only nine in Terence. Given that the Plautine text corpus is about three and a half times larger, we can see that Terence uses *faxo* only half as often as Plautus. Can we therefore conclude that *faxo* is dying out just like the other sigmatic forms? The answer is probably no. We have to take alternative causative constructions into account as well. For present purposes it is sufficient to look at the future *faciam* (first person singular). The following table presents the relevant data:

---

[25] For *ninnium* cf. Maurach (1988: 100, *ad loc.*).

	faciam (ut) +subjunctive[26]	*faciam* + double accusative[27]
Plautus	40	26
Terence	7	1

Clearly, causative constructions in general are far more frequent in Plautus than in Terence, but the ratio of *faciam ut* to *faxo* is the same for the two authors. This means that what seems to be a diachronic change in the use of *faxo* in main clauses is in reality nothing but a stylistic difference between Plautus and Terence.

3. A rationale for the distribution of the sigmatic indicative forms?

We may now summarize some of our factual findings before we try to answer the questions asked at the beginning. In main clauses only the form *faxo* appears. It is relatively frequent and shows no signs of decline in Terence. By way of contrast, we find that a number of verbs have sigmatic indicative forms in Plautus' subordinate clauses. They are moderately frequent in conditional clauses, rare in temporal and relative clauses, and non-existent elsewhere. In Terence there is only one token of these forms in a conditional clause.

Two questions arise immediately: first, why is there a variety of forms in subordinate clauses, but not in main clauses? And second, why do we only find *faxo* in main clauses, that is a first person singular indicative, and neither first persons of other verbs nor other persons of the same verb?

We need not be surprised to find a difference between subordinate and main clauses. Typologically, it is well known that these two types of clauses can obey different constraints. Quite often there are syntactic differences: in Vedic the verb is accented in subordinate clauses, but unaccented in main clauses, and in German and Dutch the word order differs in the two clause types. So the fact that some archaic high register forms exist only in subordinate clauses is not worrying. But why mainly in conditional clauses? The reason is not entirely clear, but it may well be that old and solemn legal phrases such as *si nox furtum faxit* ('if he commits larceny at night') are partly responsible. After all, legal language is especially rich in conditional clauses.

---

[26] We include the types *non faciam quin* (Plaut. *Mil*. 283) and *unum hoc faciam ut*; these all have object clauses. We exclude the adverbial *ut*-clauses in Plaut. *Persa* 662 and Plaut. *Capt*. 385-6.

[27] We include the type *faciam te proinde ac meritus es*, but we exclude the type *faciam ex te litteram longam*.

The survival of *faxo* in main clauses is only understandable if *faxo* + future or subjunctive has a special idiomatic status, either as a causative construction or as a marker of certainty (be it verbal or not). As some uses of *faxo* are clearly verbal, we must ask why we have no evidence for *faxis* or other persons of this verb in main clauses. Here there may be a series of concomitant reasons. The form appears mainly in promises and threats, and speech acts like these naturally call for first person verbs. Equally important is perhaps the fact that the type *fac* or *facito* + subjunctive (*fac habeant*, etc.) is very frequent in Plautus; its function would partly overlap with that of a causative second person *faxis* + subjunctive, so that the latter construction is superfluous.

4. Conclusions

We asked initially what the function of the sigmatic indicative forms was and whether they indicated tense or aspect. The forms do mark futurity, but there seems to be a clear split between the forms used in subordinate and main clauses. The ones in subordinate clauses alternate with the future perfect, while the isolated *faxo* in main clauses cannot indicate anteriority as the future perfects do, and seems much closer to the simple future. This naturally raises the question of how the split may have occurred and what the tokens in main and subordinate clauses have in common. The isolated *faxo* is more likely to have preserved the original value, and therefore the anterior meaning of the other forms seems to be secondary. Two factors may have caused them to take on the anteriority meaning of the future perfects: (a) the preponderance of anterior statements in some types of subordinate clauses in general and in the protasis of conditionals in particular, and (b) the fact that the sigmatic indicatives are restricted to telic verbs, and that the forms can always be interpreted as perfective.

As we have seen, the productivity of the sigmatic forms is limited. In conditional clauses, they are still a real presence in Plautus, but they are almost entirely absent from Terence. After that they only survive as archaisms. This is not surprising since their semantic content does not differ from that of the future perfect, which is much better integrated into the regular Latin tense system. To judge from Terence, *faxo* had a longer life, partly because its idiomatic uses guaranteed its survival. However, as a causative its role was taken over by *faciam ut*, while there were other contenders for its function as a marker of certainty. Morphologically, *faxo* was isolated, and that furthered its disappearance.

With regard to the question of why the forms die out in the way they do, it seems that the lack of integration into a paradigm combined with the lack of a special meaning is one reason. Moreover, it is the forms in subordinate clauses that die out first, presumably because they are stylistically marked.

There remains the question of the historical origin of the forms,[28] but that will have to be discussed in the context on the one hand of the results mentioned here, and on the other hand of the sigmatic subjunctives (*faxim*, etc.).

Appendix: the sigmatic indicative forms

Except for the first person singular, the indicative and subjunctive forms look the same. Metre can in some instances help us to disambiguate the forms, but for the most part we have to rely on syntactic and semantic criteria.

The following sigmatic forms are in the indicative:

(a) Plautus:
- conditional clauses: *adcurassis* (*Persa* 393, *Pseud.* 939a), *adempsit* (*Epid.* 363), *amasso* (*Cas.* 1001), *aspexit* (*Asin.* 770), *capso* (*Bacch.* 712), *capsimus* (*Rud.* 304), *cenassit* (*Stich.* 192), *commonstrasso* (*Epid.* 441), *comparassit* (*Epid.* 122), *decollassit* (*Cas.* 307), *ecfexis* (*Cas.* 708-12, *Poen.* 428), *enicasso* (*Most.* 212, *Most.* 223), *exoculassitis* (*Rud.* 731), *faxo* (*Bacch.* 848, *Men.* 157, *Fretum*), *faxis* (*Asin.* 612, *Capt.* 124, *Capt.* 695, *Men.* 113, *Mil.* 1417,[29] *Pseud.* 533, *Rud.* 1118, *Stich.* 610), *indicasso* (*Poen.* 888),[30] *intrassis* (*Men.* 416), *invitassitis* (*Rud.* 811), *irritassis* (*Amph.* 454, *Persa* 828, *Stich.* 345), *liberasso* (*Most.* 223), *mulcassitis* (*Mil.* 163), *occepso* (*Amph.* 673, *Cas.* 1001), *occepsit* (*Asin.* 794), *occeptassit* (*Rud.* 776), *occupassit* (*Asin.* 818), *orassis* (*Epid.* 728), *peccasso* (*Rud.* 1348), *peccassis* (*Rud.* 1150, *Rud.* 1345, *Stich.* 727), *reconciliasso* (*Capt.* 576), *respexis* (*Aul.* 58, *Rud.* 679), *servasso* (*Most.* 228)
- relative clauses: *demutassit* (*Stich.* 725), *faxit* (*Cas.* 1016)
- temporal clauses: *amasso* (*Cas.* 1002), *iniexit* (*Persa* 70), *peccassit* (*Cas.* 825)

---

[28] Sjögren (1906: 1) believes that they go back to aorists, but more recently Meiser (1998: 182-3) has suggested a desiderative origin.

[29] This form has also been classified as a subjunctive for metrical reasons. I count it as an indicative. Metre does not really help us because the final syllable is *anceps* (there is a change of speaker).

[30] This form occurs in a modifying *nisi*-clause that is semantically on a par with the preceding main clause. It does not indicate anteriority and is presumably an innovation. However, it would probably not be possible to use a non-anterior indicative form other than *faxo* in a proper main clause.

• main clauses: *faxo* (*Amph.* 355, *Amph.* 589, *Amph.* 972, *Amph.* 997, *Asin.* 132, *Asin.* 749, *Asin.* 876, *Asin.* 897, *Aul.* 578, *Bacch.* 506, *Bacch.* 715, *Bacch.* 831, *Bacch.* 864, *Capt.* 801, *Capt.* 1010, *Cas.* 484, *Curc.* 587, *Epid.* 156, *Epid.* 469, *Epid.* 656, *Epid.* 712, *Men.* 113, *Men.* 326, *Men.* 468, *Men.* 521, *Men.* 539-40, *Men.* 562, *Men.* 644, *Men.* 661, *Men.* 791, *Men.* 950, *Men.* 956, *Mil.* 463, *Mil.* 1367, *Most.* 68, *Most.* 1133, *Persa* 161, *Persa* 195, *Persa* 439, *Persa* 446, *Poen.* 162, *Poen.* 173, *Poen.* 346, *Poen.* 371, *Poen.* 460, *Poen.* 908, *Poen.* 910, *Poen.* 1154, *Poen.* 1191, *Poen.* 1227, *Poen.* 1228, *Pseud.* 49, *Pseud.* 387, *Pseud.* 393, *Pseud.* 766, *Pseud.* 949, *Pseud.* 1039, *Pseud.* 1043, *Pseud.* 1328, *Rud.* 365, *Rud.* 578, *Rud.* 800, *Rud.* 1351, *Trin.* 60, *Trin.* 62, *Trin.* 882, *Truc.* 118, *Truc.* 428, *Truc.* 643, *Truc.* 761)

(b) Terence:
• conditional clause: *faxis* (*Andr.* 753)
• main clauses: *faxo* (*Ad.* 209, *Ad.* 847, *Andr.* 854, *Eun.* 285, *Eun.* 663, *Haut.* 341, *Phorm.* 308, *Phorm.* 1028, *Phorm.* 1055)

References

Gaffiot, Félix (1933). 'Le futur antérieur dans les propositions indépendantes et principales'. *Revue de Philologie*, 59: 165-84.

Haffter, Heinz (1934). *Untersuchungen zur altlateinischen Dichtersprache.* Berlin.

Happ, Heinz (1967). 'Die lateinische Umgangssprache und die Kunstsprache des Plautus'. *Glotta*, 45: 60-104.

Kauer, Robert, and Lindsay, Wallace M. (1957). *P. Terenti Afri Comoediae.* Oxford.

Krebs, Johann Ph., and Schmalz, Joseph H. (1905). *Antibarbarus der lateinischen Sprache. Nebst einem kurzen Abriss der Geschichte der lateinischen Sprache und Vorbemerkungen über reine Latinität* (vol. 2, 7th rev. edn.). Basel.

Leumann, Manu (1977). *Lateinische Laut- und Formenlehre.* Munich.

Lindsay, Wallace M. (ed.) (1991). *T. Macci Plauti Comoediae.* Oxford (reprint of the edn. Oxford 1904/5).

Marx, Friedrich (1959). *Plautus Rudens: Text und Kommentar.* Amsterdam.

Maurach, Gregor (1988). *Der Poenulus des Plautus.* Heidelberg.

Meiser, Gerhard (1998). *Historische Laut- und Formenlehre der lateinischen Sprache.* Darmstadt.

Risselada, Rodie (2000). 'The end of time: a perfect future in Latin?', in Caroline H. M. Kroon and Daan den Hengst (eds.), *Ultima Aetas: Time, Tense*

and Transience in the Ancient World. Studies in Honour of Jan den Boeft.
Amsterdam, 25-39.

Sjögren, Hakan (1906). *Zum Gebrauch des Futurums im Altlateinischen.*
Uppsala/Leipzig.

Wessner, Paul (ed.) (1902-8). *Aeli Donati quod fertur Commentum Terenti,
accedunt Eugraphi Commentum et Scholia Bembina.* Leipzig.

On the prosody of Latin enclitics

Philomen Probert[*]

1. Introduction

When a Latin word is followed by the enclitic *-que* 'and', does *-que* cause the accent of the preceding word to appear on its final syllable? If so, does this happen only if the relevant syllable is heavy after the addition of *-que* (*laurúsque*), or also if that syllable is light (*līmináque*)? If only in the type *laurusque*, what happens in the type *līminaque*? Is the accent of *lĭmina* simply retained (*lĭminaque*), or does the accent fall on the antepenultimate syllable of the whole complex (*līmínaque*)? Perhaps the original accent is retained with the *addition* of an accent on the final syllable (*lĭmináque*). But if so, perhaps the additional accent is a secondary accent subordinated to the basic accent of *lĭmina* (*lĭminàque*). Or perhaps, on the other hand, it is the basic accent that becomes subordinated to the additional accent (*lĭmináque*). Or maybe an additional accent appears in such cases not on the final syllable of *līmina* but on *-que* (*lĭminaqué*).

All of the views mentioned above have been held, and there is at present no agreement on the subject of the accentual behaviour of *-que* or other enclitics in Latin.[1] One thing, however, is more or less generally agreed: that the Latin grammarians held that enclitics caused the accent of the preceding word to fall on the final syllable (*laurúsque, līmináque*). Some have regarded the grammarians as somewhat inconsistent on the point, but the main difficulty has been whether or not to trust the grammarians here.

The purpose of this essay is to introduce into the discussion a minor observation which may tip the balance of evidence in favour of the grammarians' view. First, however, it will be useful to survey the statements of the grammarians, the reasons which have been adduced against confidence in their statements, and the positive evidence which has been found in favour of other accounts of Latin enclisis.

---

[1] For a survey of views found in Latin grammars, see Tucker (1965: 449–50).

2. *Latin grammarians on enclisis*

When one surveys the treatments of enclisis found in the Latin grammarians one finds, firstly, some quite unequivocal statements, including an extensive discussion by Pompeius, to the effect that certain particles cause the accent of the preceding word to fall on the final syllable of that word, regardless of quantity:

(1) Servius on Vergil, *Aeneid* 10.668:
moria, id est minores particulae, ut '-que, ne, ve, ce', quotiens iunguntur aliis partibus, ante se accentum faciunt, qualislibet sit syllaba quae praecedit, sive brevis sive longa, ut 'musaque, huiusve, illucce, tantone',...

'For the *moria*, that is to say the smaller particles, like *-que, ne, ve, ce*, whenever they are joined to other words, put the accent before themselves, whatever sort of syllable precedes, whether short or long, as in *mūsáque, huiúsve, illúcce, tantṓne*.'

(2) Servius, *Commentarius in artem Donati*, Keil 4.427.6–10:
quattuor sunt particulae, quae corrumpunt in pronuntiando regulas accentuum hae, ve ne que ce. nam quotienscumque istae particulae sequuntur, faciunt accentus in ultimis syllabis superiorum esse sermonum, ut Musaque Musane Musave illiusce huiusce.

'There are four particles which disturb the rules of accents in pronunciation, namely *ve, ne, que, ce*. For whenever these particles follow, they put the accents on the last syllables of the preceding words, as in *Mūsáque, Mūsáne, Mūsáve, illiúsce, huiúsce*.'

(3) Pompeius, *Commentum artis Donati*, Keil 5.131.16–38:
pronuntiatio frequenter corrumpit (sc. *regulas accentuum), ut puta doctus. quae syllaba habet accentum? doc, quoniam prior syllaba semper habet accentum. ultima enim numquam habet aut in versu aut in prosa. adde illi coniunctionem, et vides quoniam prior syllaba perdet accentum, puta 'doctusque Palaemon'. iam incipit non in doc esse accentus, sed in tus; ecce ultima syllaba habebit accentum. et quae sunt partes quae additae corrumpunt regulas? non enim omnes corrumpunt, sed sunt partes quae additae corrumpunt. et sunt istae: ne, quando est coniunctio, ecce habes unam; ve, ecce habes duas; que, ecce habes tres; ce, ecce habes quattuor. istae sunt quae corrumpunt regulas accentuum. puta primus, adde omnes istas partes: primúsne, ecce non habebit iam accentum pri, sed mus, ecce corrumpit; et si dicas primúsve, ecce corrumpit; primúsque, primúsce, ecce corrumpit. sed forte dicas mihi 'sed positio corrumpit, non accessus istarum partium'. etiam si positio non accederet, tamen corrumperetur accentus. puta Musa, mu habet accentum; adde illi istas partes, et corrumpitur, Musáne Musáve Musáque*

Musáce: ecce illa quae est brevis ultima ipsa habet accentum. et diximus quia, quando paenultima brevis fuerit, tunc tertia a fine habebit accentum; et tamen ecce corruptam invenimus. ecce vides, in quantum valeant istae particulae, in quantum possint ut, etiam brevis si sit, id est syllaba quae naturaliter brevis est cogatur habere accentum, non naturaliter, sed causa istarum. nam ecce Musa, illa syllaba quae prior est ipsa habet accentum. si addas que aut ve, si dicas Musaque Musave, ubiubi contigerit, ipsa ultima *habet accentum.*

'Pronunciation frequently disrupts (*sc.* the rules of accents). Consider *doctus*. Which syllable has the accent? *doc*, because the first of two syllables always has the accent. For the final syllable never has the accent, whether in verse or in prose. Add to that word a conjunction, and you see that the first syllable will lose its accent: consider *doctusque Palaemon*. Now the accent begins not to be on the syllable *doc*, but on *tus*, and thus the last syllable will have the accent. And what are the words which, when added, disrupt the rules? For they do not all disrupt the rules, but there are words which disrupt them when added. And they are the following: *-ne*, when it is a conjunction — there you have one; *-ve* — there you have two; *-que* — there you have three; *-ce* — there you have four. These are the particles which disrupt the rules of accents. Consider *prīmus*, and add all these words: *prīmúsne* — there the *prī* won't have the accent anymore, but *mus*, so it disrupts the rule; and if you were to say *prīmúsve*, there it disrupts the rule; *prīmúsque, prīmúsce* — there it disrupts. But perhaps you would say to me, 'But it's position (i.e. a heavy syllable), not the addition of these particles, that disrupts the rule.' Even if position didn't arise, the accent would still be disrupted. Consider *Mūsa*: the syllable *Mū* has the accent. Add to it these words, and it is disrupted: *Mūsáne Mūsáve Mūsáque Mūsáce* — there the final syllable, which is light, has the accent on itself. And we have said that when the penultimate syllable is light the antepenultimate will have the accent; and yet here we find the rule disrupted. So you see the strength of these particles, how they can cause a syllable, even if it is light — that is to say, a syllable which is naturally light —, to have the accent, not by its nature, but as a result of these particles. For consider *Mūsa*: the first of the two syllables has the accent on itself. If you were to add *-que* or *-ve*, if you were to say *Mūsaque Mūsave*, wherever it occurs, the *last* syllable has the accent.'

Diomedes, too, offers the same rule, although the passage has occasionally been otherwise interpreted, and certain points will need clarification:

(4) Diomedes, *Ars Grammatica*, Keil 1.433.19–29:

item coniunctio complexiva sive copulativa que et disiunctiva ve et relativa[2] ne adiunctae
verbis et ipsae amittunt fastigium et verbi antecedentis longius positum cacumen adducunt
ac iuxta se proxime conlocant sic, que, ut

> *liminaque laurusque dei;*

item ve, ut

> *Hyrcanisve Arabisve putant*

et

> *calathisve Minervae;*

ne, ut

> *hominesne feraene.*

'And the connecting or copulative conjunction *-que*, the disjunctive particle *-ve*, and the relative particle *-ne* when added to words both lose their accent themselves and draw to themselves the accent of the preceding word, which was situated further away, and place it right next to themselves, as in

(*-que*)   *liminaque laurusque dei* (Vergil, *Aen.* 3.91);

(*-ve*)   *Hyrcanisve Arabisve putant* (Vergil, *Aen.* 7.605)

       calathisve Minervae (Vergil, *Aen.* 7.805);

(*-ne*)   *hominesne feraene* (Vergil, *Aen.* 1.308).'

Tucker (1965: 45, with n. 14) misunderstands the phrase *verbi antecedentis longius positum cacumen* as referring to a syllable 'long by position' and concludes that Diomedes prescribed 'presumably *Mūsámque* but *Mŭ̄saque*'. It is clear, however, that *cacumen* (or the variant reading *acumen*), like *fastigium*, refers to the accent[3] and that there is a contrast between the accent having been *longius positum* 'situated further away' and being drawn by the particles *iuxta se proxime* 'right next to themselves'.[4]

---

[2] Caesarius' conjecture *dubitatiua* would make more sense here (and is read by Corssen 1858–9: 2.257), but it is difficult to see how the corruption would have occurred.

[3] Cf. also Martianus Capella 3.273: *accentus partim fastigia vocamus, quod litterarum capitibus apponantur, partim cacumina, tonos vel sonos, Graeci prosodias* 'we call accents sometimes *fastigia*, because they are placed onto the tops of letters, sometimes *cacumina*, *toni*, or *soni*, and the Greeks call them *prosodiae*' (a passage for which *acumina* rather than *cacumina* has been conjectured, but the term *cacumen* is supported by its occurrence both here and in our passage of Diomedes; cf. Willis' apparatus for Martianus Capella 3.273).

[4] Note also the verbal parallelism of Audax, *de Scauri et Palladii libris excerpta*, Keil 7.361.5–7: *illam novitatem habent, ut et ipsae fastigium perdant et illarum partium, quibus subiunctae sunt, levationem in novissimas syllabas transferant.* 'They have the peculiarity that they both lose their accent themselves and transfer prominence to the final syllables of those words to which they are adjoined.'

Soubiran (1966: 464 n. 4) reports a personal communication from Perret taking *iuxta se proxime* to mean 'aussi près qu'il est possible (compte tenu des lois générales de l'accentuation latine)'. This is, however, also a forced interpretation. 'As near as possible' would have been *quam proxime*. Diomedes, therefore, gives the same rule that we find in Servius and Pompeius.

A passage of Probus has also been interpreted as implying a divergent rule. In discussing the syntactic functions of various particles, Probus makes a distinction between *ne* 'with grave accent' and *ne* 'with acute accent':

> (5) Probus, *Instituta artium*, Keil 4.145.21–5:
> *item ne, si gravem accentum habeat, erit coniunctio, ut puta*
> > *tantane vos generis tenuit fiducia vestri;*
> *si vero ne acutum accentum habeat, erit adverbium, ut puta*
> > *ne credite, Teucri,*
> > *quidquid id est.*

> 'And *ne*, if it has a grave accent, will be the conjunction, as in
> > *tantane vos generis tenuit fiducia vestri?* (Vergil, *Aen.* 1.132);
> but if *ne* has the acute accent, it will be the adverb, as in
> > *ne credite, Teucri.*
> > *quidquid id est* (Vergil, *Aen.* 2.48–9).'

Wagener (1904: 506) takes this passage to imply that Probus said *tántane*. Kent (1945: 68 n. 1) similarly cites Probus here as giving an opinion 'dissenting' from the usual prescription by Roman grammarians of *armáque*, *generáque*, etc. However, nothing in our passage implies an accentuation *tántane* as opposed to *tantáne*. Probus is contrasting postposed, enclitic -*ne* with non-enclitic *nē*. The characterisation of enclitic -*ne* as having a grave accent implies only that enclitic -*ne* was unaccented. Non-enclitic *nē*, on the other hand, had the 'acute accent', which is to say that it was accented. Probus does not tell us whether the accent of *tantane* fell on the first or the second syllable, only that it did not fall on the third. Wagener appears to regard the use of Vergil, *Aen.* 1.132 as implying the accentuation *tántane*, the point presumably being that in this line the ictus falls on the first syllable of the word. That such reasoning is mistaken is, however, clear from the line of Vergil used to exemplify accented *nē* (*Aen.* 2.48), in which the verse ictus does not fall on *nē*. Ictus and accent do not necessarily coincide in Classical verse, and we cannot assume that lines of verse cited by grammarians in discussions of accent are chosen for coincidence between ictus and accent: they are cited simply to exemplify the forms under discussion.

Probus' discussion is therefore entirely consistent with the rule we have already encountered, which would suggest that *tantane* was in fact accented *tantáne*.

The same rule is given, though without explicit mention of the case where the syllable immediately preceding the enclitic is light, by Martianus Capella (3.272) — whose ultimate source has been plausibly identified as Varro —,[5] Priscian (*Partitiones*, Keil 3.466.2–5, 3.477.1–3), and Audax (*de Scauri et Palladii libris excerpta*, Keil 7.361.2–9). In all these passages the particles said to be relevant for the rule are *que, ve*, and *ne*; *ce* occurs (as in the passages from Servius and Pompeius quoted above) in some of the manuscripts of Martianus Capella.

The passage of Martianus Capella specifies that, in order for *-que* to be relevant to the rule, it must be the *complexiva coniunctio* 'connecting conjunction', while *ve* must be, rather obscurely, the *expletiva* (*coniunctio*):

(6) Martianus Capella 3.272:

mutant accentus adiunctis vocibus que, ve, ne, *cum tamen complexiva coniunctio est* que, ve *[cum] expletiva, ut* Latiumque augescere vultis et stimulove meum cor *apud Accium in Pelopidis. numquam migrabit acutus sonus de primis syllabis in postremas praeter particulas coniunctas, quarum hoc proprium est acuere partes extremas vocum, quibus adiunguntur.*

'Accents change when the words *que, ve, ne* are adjoined, as long as *-que* is the connecting conjunction, *ve* the expletive conjunction, as in *Latiumque augescere vultis* (Ennius, *Annales* 455 V.) and *stimulove meum cor* in Accius' *Pelops* (Accius, *trag.* 512). The acute accent will never move from a first syllable to a last except in the case of adjoined particles, since it is peculiar to them to cause an acute accent on the last part of the words to which they are adjoined.'

[Priscian]'s *De accentibus liber* provides an example of a word with non-connective *-que* that does not come under the enclitic rule, the conjunction *ítaque* 'therefore', as contrasted with *itáque* 'and thus', in which *-que* is the connective:

---

[5] So Schöll (1876: 7), pointing out that the words *mutant accentus* used here by Martianus recall Gellius' remark (18.12.8) that Varro used the phrase *mutant accentus* instead of *mutantur*. Schöll notes that intransitive use of the active *muto* is attested three times in Varro's *De lingua latina* (5.101, 5.170, 9.64, plus a conjecture at 9.54) in addition to the attestation preserved by Gellius. Schöll also observes that in the relevant passage Martianus quotes lines of Ennius and of Accius, authors with whom he was not directly acquainted.

(7) [Priscian], *De accentibus liber*, Keil 3.520.36–3.521.4:
necessitas pronuntiationis regulam corrumpit, ut puta siquis dicat in primis dóctus, addat que coniunctionem dicatque doctúsque, ecce in pronuntiatione accentum mutavit, cum non in secunda syllaba sed in prima accentum habere debuit. sunt quidem syllabae tres, in quibus accentus producitur, que ne ve, ut itáque, quando adverbium est; quando vero coniunctio, ítaque dicimus: 'venerúntne viri ad nos'? 'carbonibúsve lumbos assarunt'.

'Exigencies of pronunciation disrupt the rule, as when somebody to begin with says *dóctus*, then adds the conjunction -*que* and says *doctúsque* — so he's changed the accent in pronunciation, since the word ought to have had the accent not on the second syllable but on the first. There are three syllables which cause the accent to move forward — *que ne ve* —, as in *itáque*, when *ita* is an adverb. But when the word is a conjunction, we say *ítaque*: *venerúntne viri ad nos*? *carbonibúsve lumbos assarunt*.'

Priscian implies in passing the same distinction between *ítaque* 'therefore' and *itáque* 'and thus' when he says:

(8) Priscian, *Institutiones*, Keil 3.100.15–16:
Collectivae vel rationales sunt 'ergo, igitur, itaque', quando antepaenultima acuitur, 'quin, alioquin, immo, utique, atqui'.

'The syllogistic or rational conjunctions are *ergo*, *igitur*, *itaque* (when the antepenultimate is accented), *quin*, *alioquin*, *immo*, *utique*, *atqui*.'

The example of *ítaque* reappears, with three others, *útique* 'certainly' (as contrasted with *utíque* 'and that'), *úndique*, and *dénique*, in two further passages of Priscian. We consider first the fuller discussion, in the *Institutiones grammaticae*. Here, while discussing feminine forms of adjectives whose declensions are influenced by the corresponding masculines, Priscian claims that influence from masculine to feminine can also affect accentuation:

(9) Priscian, *Institutiones grammaticae*, Keil 2.181.16–2.182.3:
nec solum tamen in declinationibus nominum hoc contigit, sed etiam in accentibus, ut 'utérque utráque utrúmque', 'plerúsque pleráque plerúmque': femininum enim, quamvis paenultima brevis sit, accentum tamen in ea habuit acutum, sicut masculinum et neutrum: 'que' enim, nisi separata sit, si Graecos sequimur, coniunctio enclitica esse non debet, nisi illud dicamus, quod 'que', quando cum integris componitur dictionibus, quamvis significationem suam amittat [id est coniunctionis], tamen enclitici vim servat: 'pleráque', 'ubíque', 'utráque', exceptis differentiae causa 'ítaque', 'útique'; in his enim non solum coniunctio, sed etiam praeposita ei adverbia vim propriae significationis convertunt

*composita. vetustissimi tamen et 'altera utra' et 'alterum utrum' et 'alterius utrius'
solebant proferre et 'plerus plera plerum' absque 'que' additione.*

'It is not, however, only in the declensions of nouns that this happens, but also in accents, as in *utérque utráque utrúmque*, *plerúsque pleráque plerúmque*: for the feminine, although it has a light penultimate, nevertheless has the acute accent on it, like the masculine and neuter. For *que*, if we follow the Greeks, ought not to be an enclitic conjunction unless it be a distinct word — unless we say that *que* keeps the force of an enclitic when it is compounded with whole words, even if it loses its meaning [that of a conjunction]: *pleráque*, *ubíque*, *utráque*, with the exception of *ítaque* and *útique*, which are accented as they are for the sake of differentiation (i.e. from *itáque* 'and thus' and *utíque* 'and that'). For in these not only the conjunction but also the adverbs that are placed before it lose the force of their particular meaning when compounded. But the ancients used to say *altera utra*, *alterum utrum*, *alterius utrius*, and *plerus*, *plera*, *plerum*, without the addition of *que*.'

Allen (1989: 88), not accepting the rule that a word followed by an enclitic was accented on its final syllable regardless of quantity, suggests that there may, nevertheless, have been accentuations of type *bonáque* following forms in the same paradigm such as *bonúsque*. He adduces our passage of Priscian in support of such analogical accentuations in the case of *utráque* and *plēráque*, but reserves judgement on the applicability of such accentuations to the Classical period.

It is important to note that Priscian's passage, taken as a whole, does not conflict with the enclitic rule we have already seen but in fact presupposes it. The accentuation of *utráque* and *plēráque* is indeed taken, in the first instance, to be due to the influence of the masculines *utérque* and *plērúsque* and the neuters *utrúmque* and *plērúmque*. The following discussion of the circumstances under which *-que* is an enclitic implies, however, that if *-que* were taken to be an enclitic in *utraque* and *plēraque*, the principles of enclitic accentuation would account for the accentuations *utráque* and *plēráque*. The difficulty appears to be whether to take *-que* to be an enclitic only where it retains its own meaning ('and') or to allow that it is an enclitic wherever it is added to a whole word. It emerges that *itaque* 'therefore' and *utique* 'certainly', where *-que* does not mean 'and', are accented on their initial syllables. This would pose no difficulty if *-que* did not count as an enclitic where it did not mean 'and'. If, however, *-que* is an enclitic wherever it has been added to a whole word, then the rule for enclitics would imply the accentuations **itáque* 'therefore' and **utíque* 'certainly', so some exception would need to be made for *ítaque* and *útique*.

Priscian suggests two possible reasons — not mutually exclusive — why *ítaque* and *útique* might be exceptions to the enclitic rule even if they ought to

fall under it, i.e. if *-que* is to be taken as an enclitic whenever added to a full word. The first reason, *differentiae causa*, appeals to the fact that *itáque* 'and thus' and *utíque* 'and that', where *-que* means 'and', also exist; *ítaque* 'therefore' and *útique* 'certainly' would be accented on the first syllable to differentiate them from *itáque* 'and thus' and *utíque* 'and that'. Priscian's second suggestion is that in *ítaque* 'therefore' and *útique* 'certainly' the first element as well as the second has lost its proper meaning.

Something should be said about the quantity of the vowel in the penultimate syllable of *ítaque, útique, itáque,* and *utíque*, as well as in *ubíque* 'everywhere', which Priscian also adduces as an example of *-que* not meaning 'and' but being added to a whole word. There is no doubt that *ítaque* 'therefore' and *itáque* 'and thus' have a short *-a-* in the penultimate syllable. Both words occur in poetry, especially in Plautus and Terence, always with metrically light penultimate syllable except at *Arg.* Plaut. *Cist.* 10 (cf. the *Oxford Latin Dictionary*, s.v.). The quantity of the vowel in the second syllable of *útique* 'certainly' and *utíque* 'and that' is more difficult. As far as I have been able to ascertain, the words do not occur in verse. Priscian's discussion presupposes a short vowel for both words, but an early form of *utique* 'certainly' is attested with the spelling *VTEIQUE* (CIL 1.585.79), which should imply a long vowel for the Classical form. Apparently making different judgements on the basis of this inadequate evidence, Lewis and Short in their *Latin Dictionary* indicate a short *-i-* for *utique* 'certainly', while the *Oxford Latin Dictionary* indicates a long *-ī-*. The simplex form *utī* 'that' has a long *-ī*, and this should make it likely that *utíque* 'and that' also had a long *-ī-*, since this form is a transparent concatenation of *utī* and *-que*. The form *ubíque* when it occurs in poetry always has a metrically heavy penultimate syllable, implying a long *-ī-* (see the *Oxford Latin Dictionary*, s.v.). Priscian thus discusses various words as if they all had a light penultimate syllable, although it is likely that this was the case for only some of them, most clearly for *ítaque* and *itáque*, as well as *plēráque* and *utráque*. Since the contrast between long and short vowels had been lost by Priscian's time, it is perhaps not surprising that there is some confusion here.[6] Since, however, for our purposes it

---

[6] However, the vowels that in Classical Latin had been *ī* and *i* remained distinct in *quality* even when the difference in *quantity* had been lost. I suspect that Priscian's assumption that *útique*, *utíque* and *ubíque* all have a short *-i-* makes it likely that all of these words had at least variant forms with short *-i-*, and that these gave rise to the forms known to Priscian. Long *ā* and short *a*, on the other hand, did not differ in quality and therefore simply fell together when quantitative distinctions were lost. This explains the otherwise curious fact that Servius (*Commentarius in artem Donati*, Keil 4.427.13–19) treats the phonological difference between *ítaque* 'therefore' and *itáque* 'and thus' as a matter of the vowel in the penultimate syllable being short in the first case, long in the second. For Classical Latin, we know that the *-a-* was short in both cases. For Servius in the fourth century AD, short and long *-a-* were no

is the words which in the Classical period had a light penultimate syllable whose accentuation is particularly at issue, we need to keep in mind that only some of the words Priscian discusses here clearly fall into this category.

To return to Priscian, he takes the following points for granted. Firstly, there is a general rule according to which an enclitic causes an accent to fall on the final syllable of the preceding word. Secondly, *utráque*, *plēráque*, *ubíque*, *itáque* 'and thus', and *utíque* 'and that' are accented on the syllable immediately preceding *-que*, while *ítaque* 'therefore' and *útique* 'certainly' are not. Of the words which, according to Priscian, have the accent on the final syllable of the word preceding the enclitic, *utráque*, *plēráque*, and *itáque* 'and thus' certainly have a light penultimate syllable and are therefore relevant to our question. Of the words which Priscian assumes to have the accent on the antepenultimate, *ítaque* 'therefore' again certainly has a light penultimate syllable. An antepenultimate accent on *útique* 'certainly' would be completely unexpected if the penultimate syllable were heavy, but there is insufficient evidence to verify the quantity of the *-i-*, and therefore the weight of the penultimate syllable.

Priscian is, importantly, somewhat embarrassed by the rather messy set of facts he presents: he cannot define the circumstances under which *-que* is an enclitic in such a way as to account neatly, without need of further pleading, for the forms in which an accent falls on the syllable immediately preceding *-que* and the small number of forms in which it does not. Priscian's embarrassment here should give us some confidence that the facts themselves are not simply grammarians' inventions: he does not have a neat theory into which his facts conveniently fall and on the basis of which they could easily have been invented.

Priscian's less extensive discussion of the same facts, in the *Partitiones*, suggests only the explanation that *utráque* and *plēráque* are accented by analogy with the corresponding masculines and neuters:

---

longer distinct, but application of the normal rule of accentuation for Classical Latin words (disregarding the question of enclitics) would have given rise to the accentuation *ítaque* if the *-a-* was short, *itáque* if the *-a-* was long. Servius thus extrapolates incorrectly from the forms known to him the Classical quantities of the vowels in the penultimate syllables of these forms. His statement provides further evidence for an accentual difference between *ítaque* 'therefore' and *itáque* 'and thus', since this difference is necessary to explain his error concerning the vowel quantities.

(10) Priscian, *Partitiones*, Keil 3.488.24–30:

syllabice quoque additur (sc. -que) in fine et vim coniunctionis amittit encliticae, ut undique denique itaque, quando antepaenultimo acuto profertur, id est quando coniunctio est rationalis. utraque et pleraque, quaeritur, cum sint una pars orationis nec que divisa pro coniunctione accipiatur, cur non tertium ab ultima acutum habuerunt. in quo possumus dicere, quod accentus masculinorum et neutrorum, quia paenultimus est acutus, acuit etiam feminina in suam legem concidentia.

'It (*sc. -que*) is also added at the end of a word as a syllabic extension and loses the force of an enclitic conjunction, as in *undique, denique, itaque* (when it is pronounced with an acute on the antepenultimate — that is when it is a rational conjunction). It is a question why *utraque* and *pleraque* do not have the acute on the antepenultimate, although they are single words and *-que* is not taken separately as a conjunction. On this we can say that the accent of the masculines and neuters, since it is an acute on the penultimate, puts an acute also on the feminines so that they fall under its law.'

This passage needs to be taken in the light of the longer passage (9) quoted above. The difficulties spelled out in the longer passage make it impossible to take the shorter passage as evidence against the general rule that enclitics cause an accent to appear on the preceding syllable. Rather, as in the longer passage, the general rule is presupposed, but a non-enclitic *-que* is supposed to be present in *undique, dēnique* and *itaque* 'therefore'. On the other hand, it is apparently this same *-que* that is preceded by an accented syllable in *utráque* and *plēráque*. But the accents of these two words can be explained in another way, namely by appeal to the corresponding masculines and neuters. Again the presuppositions are that there is a general rule by which enclitics cause the preceding word to be accented on its final syllable, regardless of quantity, that *utráque* and *plēráque* are accented on the syllable preceding *-que*, as is *itáque* 'and thus', but that *úndique, dénique* and *ítaque* 'therefore' are accented on their initial syllables.

We conclude our survey of Latin grammarians' statements on enclitics with the following passage of Priscian, which provides information on a point we have not yet touched. Care is again necessary over interpretation:

(11) Priscian, *Partitiones*, Keil 3.488.19–23:

est autem encliticum et haec (sc. -que) et aliae duae coniunctiones, ve et ne. ne autem solet etiam abiecta e encliticam vim possidere, ut 'Pyrrin, tanton'. quae enclitica, id est que ve ne, praepositionibus ante se positis non erigunt extremae syllabae fastigium, ut 'propterque illum', 'interve homines'.

'This conjunction (*sc. -que*) and two others, *-ve* and *-ne*, are enclitics. And *-ne* tends to have enclitic force even when the *-e* has been lost, as in *Pyrrīn, tantōn*. These enclitics — *que*, *ve*, *ne* — do not raise up the accent of the last syllable when prepositions have been placed in front of them, as in *propterque illum, interve homines.*'

Tucker (1965: 451, with n. 15) regards Priscian as saying that 'the accent is not shifted when an enclitic follows a preposition, as *própterque illum, ínterve homines*'. The phrase *(non) erigunt extremae syllabae fastigium*, however, imitates the Greek phrase ἐγείρει τὴν ἐπὶ τέλει αὐτῆς βαρεῖαν (Arcadius 167.1), *vel sim.*, which refers to the 'awakening' by a following enclitic of a 'sleeping' acute on a final syllable. The Greek phrase is used both of cases where the 'awakened' acute is the second accent on the word preceding the enclitic (as in ἄνθρωπός τις: Arcadius 167.2) and of oxytone words in which the acute on the final syllable is the natural accent but is 'asleep' unless followed by an enclitic or by pause (as in αὐτός μοι: Arcadius 167.14). Priscian's adoption of the language used for describing Greek acute accents on final syllables induced by following enclitics therefore tells us only that there was no accent on the final syllable of a preposition followed by an enclitic. We are not told whether there was an accent on an earlier syllable of the preposition.

We have, in fact, some evidence that both syllables of a Latin preposition were normally unaccented. Audax' excerpts from Scaurus and Palladius include a discussion of the parts of speech which are sometimes unaccented. All prepositions, with the exception of *ergō* (properly a postposition), are said to be unaccented:

(12) Audax, *de Scauri et Palladii libris excerpta*, Keil 7.360.25–7.361.2:
praepositiones vero omnes sine fastigio sunt. sane notandum est 'ergo', ut 'illius ergo venimus', quod propter distantiam coniunctionis 'ergo' in posteriore syllaba circumflectitur.

'All prepositions are without an accent. One needs to note the exception *ergō*, as in *illius ergo venimus*, which has a circumflex on the last syllable to differentiate it from the conjunction *ergō*.'

Consistently with Audax' statement, Priscian states that the preposition *sine* has grave accents on both syllables when preposed but an acute on the first syllable when postposed:

(13) Priscian, *Institutiones grammaticae*, Keil 3.52.19–25:

'sine' autem etiam verbum est imperativum, sed accentu differt, quando praeponitur, quomodo 'pone': praeposita enim gravatur utraque syllaba, ut 'sìnè timore', postposita paenultimam acuit, ut Virgilius in X:

   Si sìnè pace tua atque invito numine Teucri.
idem in III georgicon:

   Te síne nil altum mens incohat.

'*sine* is also a verb in the imperative, but there is a difference in accent when the word is preposed, in the same way as for *pone*: for when preposed it (*sc.* the preposition *sine*) has grave accents on both syllables, as in *sìnè timore*, but when postposed it has an acute on the penultimate, as in Vergil's tenth book:

   Si sìnè pace tua atque invito numine Teucri (Vergil, *Aen.* 10.31).
But in the third Georgic:

   Te síne nil altum mens incohat (Vergil, *Georgics* 3.42).'

Since our evidence for the accentuation of preposed prepositions suggests that they were unaccented, the point Priscian intends in passage (11) is most likely to be that preposed prepositions simply remain unaccented when followed by enclitics: *propterque íllum, interve hómines*. If so, an enclitic following a preposition would fail to shift the accent on the preceding word only because there was, as it were, no accent to be shifted: the sort of unemphatic prepositive that did not ordinarily carry an accent did not suddenly become emphatic and accented merely because it was followed by an enclitic.

We can now sum up our findings from the Latin grammatical tradition. We are presented with a remarkably consistent tradition that certain particles, *-que*, *-ne*, *-ve*, and *-ce*, cause the word they follow to be accented on its final syllable, regardless of syllabic quantity. Preposed prepositions, however, are unaccented and remain unaccented when followed by enclitics. The words *dénique*, *úndique*, *ítaque* 'therefore' and *útique* 'certainly' are accented on their initial syllables (the last two contrasting with *itáque* 'and thus' and *utíque* 'and that'; the quantity of the vowel in the penultimate syllable of *útique* and *utíque* is, however, uncertain). These accentuations would make it attractive to suppose that *-que* was not an enclitic when it did not mean 'and', but for the awkward data *utráque* and *plēráque*: are these influenced by their corresponding masculines, or is *-que* after all an enclitic even when it does not mean 'and'? In the latter case, how does one explain *dénique*, *úndique*, *ítaque* 'therefore' and *útique* 'certainly'? These questions remain unresolved in the tradition.

3. Modern doubts about the grammarians' rule

Most pronouncements by Latin grammarians on Latin accentuation tend to be viewed with some suspicion owing to the definitely unfortunate tendency of Latin grammarians to describe the accent of their language in terms remarkably close to those used by Greek grammarians for the Greek accent. Allen (1989: 84) says, 'It is inconceivable that Latin should have developed a system of pitch accents that agreed in such minor detail with Greek, and we can only assume that the grammarians have slavishly misapplied the Greek system to the description of Latin...'. This remark is aimed especially at the fact that the Latin grammarians not only speak as if their language has a distinction between circumflex and acute accents but reproduce exactly the Greek rules for determining whether an accent on a non-final syllable is an acute or a circumflex. But the Latin rule for enclitics also comes under suspicion of being constructed under the influence of Greek enclitic accentuation. The example *Mūsáque* (see passages (1), (2), and (3) above), in particular, is suspiciously reminiscent of Greek Μοῦσά τε (see Allen 1989: 87).

It is worth noting, however, that the Latin grammarians' rule for enclitics, unlike their rules for determining whether an accent is acute or circumflex, is far from being simply a copy of the Greek rules for enclitic accentuation. Greek enclitics do not simply shift an accent from a non-final to a final syllable on the preceding word; they may make no change to the accent of the preceding word (ποικίλος τε), or add a second accent to the word (Μοῦσά τε), or 'awaken' a 'sleeping' acute on a final syllable (αὐτός μοι). None of these corresponds to the *shifting* of the accent prescribed for *Mŭsa* in *Mūsáque*, or for *húius* in *huiúsve*. We must, therefore, allow that in the case of enclitic accentuation the Latin grammarians pay at least some attention to what actually happens in their language.

But how much attention? We might expect that when the syllable immediately preceding an enclitic is heavy the accent would fall on that syllable, as in *doctúsque*. This would arise from the normal rules of Latin accentuation, assuming that the addition of an enclitic to a word created a new word in which the enclitic provided the final syllable, and the normal rules of Latin accentuation simply applied to the whole group. This apparently easy explanation of the type *doctúsque*, combined with the observation that most of the grammarians' examples of enclitic accentuation involve a heavy syllable immediately preceding the enclitic (as in *doctúsque*), have led many scholars to assume that the grammarians' rule was only valid when the syllable immediately preceding the enclitic was heavy. If so, the grammarians' contention that the rule

applied regardless of the quantity of the syllable preceding the enclitic would be an over-generalisation (so Allen 1989: 87).

It is difficult, however, to reconcile the possibility that the grammarians' rule may be an over-generalisation with the care that some of our grammatical passages take to insist that the rule applies both when there is a light syllable and when there is a heavy syllable preceding the enclitic (see passages (1) and (3) above), especially when an accentuation such as *Mūsáque* simply is not equivalent to that of Greek Μοῦσά τε.

We are left with the problem as to whether the Latin enclitic accent rule we are given by the grammarians is unlikely on typological grounds: can we believe that an accent system behaved as our Latin grammarians tell us that theirs behaved? For many scholars, the implicit answer to this question seems to have been 'no'. While the motivation for the shift of accent from *dóctus* to *doctúsque* appears to many to be clear — the normal Latin accent rule necessitates the shift of accent —, such a view of the reason for the shift here makes it impossible to accept a similar shift from *Mǔsa* to *Mūsáque*. Hence the widespread reluctance to accept the grammarians' rule in the general form in which the grammarians give it.

Recently, however, there have been dissenting voices defending the typological plausibility of the grammarians' enclitic accent rule. Steriade (1988: 298) and Halle (1997: 302) account for Latin enclitic accentuation, following the grammarians' rule, within a general typology of accent systems and using procedures that can be paralleled from other languages. This is not the place to review their account of Latin accentuation, but it is worth drawing attention to the fact that there is now a growing body of evidence to suggest that the grammarians' accent rule is not especially implausible from a typological point of view.

4. *Positive evidence adduced in favour of alternative rules*

Some modern scholars who reject the grammarians' rule have argued for various alternative rules on the basis of metrical evidence. If one can determine metrical positions in which certain poets insist on coincidence between ictus and word accent, the occurrence of words followed by enclitics in these positions ought, arguably, to help in determining where the accent falls on such words. Suggestions along these lines have often been made, and there is not space here to discuss every assertion that has appeared, but we shall take a look at the fairly detailed arguments of Tucker (1965) and then examine a somewhat different line of argument that has been explored by Shipley.

Tucker's study appeals, for the first century AD, to the iambic trimeters of Seneca and the anonymous *Octavia*. With a small number of exceptions (for which Tucker finds various explanations), these obey the rule that 'in the third foot, the long syllable of the iambus is the accented syllable of a word; or, if it is resolved into two short syllables, the first of these is the accented syllable of a word' (Tucker 1965: 453). Where a word is followed by an enclitic, Seneca consistently treats the composite just like a normal word, with the new penultimate syllable being counted for the purposes of this rule as the accented syllable if it is a heavy syllable, the new antepenultimate syllable being counted as the accented syllable otherwise.

The difficulty with Tucker's account lies in knowing whether the Senecan constraint is genuinely a constraint on the position of the accent rather than on the occurrence of word-end. As Tucker recognises (1965: 454–5), the fact that Seneca adheres strictly to the placement of word-end either after the second anceps or after the second short makes it almost inevitable that, at least for an ordinary word (not involving an enclitic), the accent will fall on the third long, i.e. in the position envisaged by Tucker's rule. The relatively rare word-shapes for which the accent does not fall in this position do occur, and Tucker's explanations of these cases involve some circularity. Thus, he finds two examples of four-syllable words in which the first syllable falls at the third long: *machinatrix* at *Med.* 266 and *classibusque* at *Agam.* 221. In both cases (and whatever view one takes of the enclitic rule), the Classical Latin accent would have fallen on the penultimate syllable of the word, not on the first syllable.[7] Tucker suggests that 'we evidently have both a primary and a secondary accent, and perhaps either one is acceptable to satisfy the rule.' Other exceptions to his rule include four-syllable words in which the first three syllables are light and the first two occur at the resolved third long: *Danaides* at *H. F.* 757, *maleficae* at *Tro.* 752, *miserias* at *Med.* 253, *Domitio* at *Oct.* 249, and *facinore* at *Hipp.* 1186. These lead Tucker (1965: 454, with n. 24) to argue that the prehistoric Latin initial accent, which survives in words of this type in Plautus, may have survived over two centuries longer than is generally supposed, and that Cicero, *Orator* 58 overlooks words of this type when he states that the Latin accent never falls more than three syllables from the end of the word. In view of the circularity involved here, we cannot take the positioning of sequences such as *nātōsque* at *H. F.* 310 or *verbaque* at *Hipp.* 1175 as evidence for their

---

[7] In the case of *machinatrix*, the accent would have fallen on the penultimate syllable in this line because Seneca treats the *-tr-* cluster as heterosyllabic here. For the accentual consequences of the variable rhythmic treatment of stop plus liquid clusters, cf. Quintilian 1.5.28.

accentuation. The only clear factors determining their positioning are the quantitative metrical pattern and the necessity for caesura.

For the second century BC, Tucker believes that the situation was different and that enclitics simply left the accentuation of the preceding word unchanged. This conclusion is drawn from the occurrence of iambic shortening in Plautus and Terence. Iambic shortening normally affected only unaccented syllables, and Tucker finds that there are numerous instances where an originally heavy syllable is 'shortened' even though it is followed by an enclitic. However, as Tucker recognises (1965: 458), iambic shortening is subject to much analogical influence. It is quite possible, therefore, that a word which would undergo shortening when an enclitic did not follow continued to undergo shortening even when an enclitic followed, by analogy with the form without enclitic. Although we cannot know that the accentual behaviour of Latin enclitics remained constant from the archaic period to late antiquity, we should be wary of drawing conclusions about change from pieces of metrical evidence which are highly uncertain in themselves and do not support one another, even if the lack of mutual support could in theory be explained as the result of chronological change.

Shipley (1913)[8] argues that in Cicero's time combinations of full word plus enclitic were accented according to the normal Latin rules for accenting single full words (so *doctúsque, ármaque, līmínaque, éaque*). His argument is based on statistics which suggest that, in essence, Cicero in his speeches positively sought to place the enclitic *-que* in a position where the preceding syllable would be heavy (*eumque, aditumque*, etc.) and actively avoided placing *-que* in a position where the preceding syllable would be light (*multaque, uidendaque, crīminaque, positaque*). The only exception to this generalization which he finds concerns the trisyllabic type ◡◡ + *que* (*eaque, suaque*, etc.), which seems to have been avoided in Cicero's earlier speeches but preferred in the later speeches.

Shipley's explanations of these findings run as follows. The preference for combinations involving a heavy syllable before *-que* (*eumque, aditumque*, etc.) is said to be motivated by the importance of trochaic sequences in Cicero's clausulae. The type ...−◡ + *que* (*multaque, uidendaque*, etc.) is said to be avoided because Cicero's rhythmic prose avoids dactylic sequences except before a pause in sense (see Shipley 1911*a*), and a pause in sense tends not to follow *-que*. Shipley finds it impossible that the type ... −◡ + *que* would have been avoided if the accent had followed the grammarians' rule (*multáque, uidendáque*). If thus accented, such combinations would 'fit beautifully' into a number of Cicero's rhythmical forms (which Shipley takes — as have others —

---

[8] See also Shipley (1909).

to be defined by the position of the accent and not only by the positions of heavy and light syllables). Relevant rhythmical forms include ⏓⏑⏑́ (accommodating (*mul*)*táque* ⏑⏑́), ⏓⏑–⏑́ (accommodating (*mul*)*táque* –⏑́), and ⏑́–⏓ (accommodating ⏑́ *multáque*; see Shipley 1913: 31). Moreover, the avoidance of a sequence because of its dactylic rhythm would, in Shipley's view, only be credible if the sequence were indistinguishable from a 'pure dactyl' such as *cóndere* (i.e. one in which the accent fell on the first syllable). A 'pure dactyl' might suggest the fifth foot of a dactylic hexameter, with regular coincidence between ictus and accent (cf. Shipley 1911*b*).

For the type ... –⏑⏑ + *que*, either accentuation according to the grammarians' rule (*crīmináque*) or accentuation according to the normal rule for single full words (*crīmínaque*) would give forms which would fit well into standard Ciceronian clausulae (Shipley 1913: 33–4). The avoidance of this type cannot, therefore, be attributed to inappropriate rhythm, whatever view of the accentuation is taken. Shipley assumes, on the basis of his argument regarding the type *multaque* (in his view *múltaque*), that the accent was in fact *crīmínaque*, following the normal rule for single full words. He then explains the avoidance of the type *crīmínaque* as due to an awkwardness felt in the shift of accent from the syllable that normally carried the accent to the syllable which he regards as having been 'ordinarily the weakest syllable in the word' (Shipley 1913: 34).

For the type ...⏑⏑⏑ + *que* (*positaque*), also avoided by Cicero, Shipley cannot derive evidence for the accent from clausulae, since Cicero's clausulae avoid words ending in four light syllables even when the final syllable is not an enclitic (Shipley 1913: 37 n. 3). However, Shipley finds evidence in the avoidance of the type for, once again, reluctance to shift the accent from the syllable which normally carried the accent. He regards the accentuation which would result from application of the normal rule for single full words as more 'awkward' than that which would result from the grammarians' rule: 'This points to the operation of the penultima law rather than to the accentuation indicated by the grammarians, since Cicero was clearly avoiding an awkward accentuation, while, if it were the rule to accentuate the syllable which precedes the enclitic, no such awkwardness would in that case be felt' (Shipley 1913: 37).

The majority of examples of the trisyllabic type ⏑⏑ + *que* concern *eaque*, and Cicero's increased preference for the type in his later speeches is in fact due to the increasing frequency of *eaque* (Shipley 1913: 38). In the *Philippics*, there are four occurrences in clausulae, which Shipley takes as evidence for the accentuation *éaque*:

> *multa praetereo eaque praeclara.* (*Phil.* 1.3)
> *multa ex me eaque saepissime.* (*Phil.* 1.34)
> *non multa eaque maculosa.* (*Phil.* 2.73)
> *eaque defendunt.* (*Phil.* 7.2)

The general avoidance of -*que* following a light syllable in Cicero's speeches is indeed striking. Shipley's explanations of the phenomena are, however, unsatisfactory for a number of reasons.[9] His argument for *múltaque* rather than *multáque* hangs on the notion that *multáque* would 'fit beautifully' into Cicero's rhythmical prose and therefore would not be avoided. The fact that the type *is* avoided is thus taken as evidence for the accentuation *múltaque*. The type *crīmínaque*, accented as Shipley supposes, would however also 'fit beautifully' (Shipley 1913: 34), and yet in this case the conclusion is not drawn that the accentuation *crīmínaque* is impossible. Rather, the accentuation *crīmínaque* is *assumed* because of *múltaque*, and the avoidance of *crīmínaque* then put down to awkwardness of accentuation. But if the fact that *crīmínaque* would fit beautifully into Cicero's rhythmical prose does not make its avoidance rule out that accentuation, neither does the fact that *multáque* would fit beautifully make its avoidance rule out that accentuation, and with that the reason for preferring the accentuation *crīmínaque* to *crīmináque* also disappears.

The reasons for thinking that *posítaque* involves a more 'awkward' accentuation than would *positáque* are not made clear, and indeed the notion that *positáque* would involve no awkwardness if the regular rule was for the accent to fall on the syllable preceding the enclitic could surely be applied also, *mutatis mutandis*, to *posítaque*: if the rule was for a combination of full word plus enclitic to be accented like a single full word, such a procedure should not give rise to 'awkwardness' any more than a rule producing *positáque*.

Shipley takes pains to show that avoidance of the types *crīminaque* and *positaque* is not a feature only of rhythmical prose (Shipley 1913: 35–6, 37), and indeed this is what one would expect if his explanation is correct. However, a mainstay of his argument that these types are avoided also in non-rhythmical prose is their avoidance in Cicero's letters to Atticus, yet he reveals that the type *multaque* is also avoided as much in the letters to Atticus as in the speeches

---

[9] This is not the place to review the methodology which leads Shipley to his conclusions as to which combinations with -*que* occur more frequently than one would normally expect and which occur less frequently. There is a certain vagueness in Shipley's statement that his figures for the frequencies of occurrence of words of certain metrical shapes in the language as a whole 'are based upon composite statistics for a considerable number of authors and include colloquial and legal Latin' (Shipley 1913: 26), and it may be worthwhile to re-examine these figures using a clearly-defined corpus.

(1913: 32). Given his explanation of the avoidance of the type *multaque*, one would expect greater avoidance in rhythmical than in non-rhythmical prose. The lack of difference in this respect between Cicero's speeches and his letters to Atticus therefore casts serious doubt on Shipley's explanation of the avoidance of the type and hence, again, on his contention that the accentuation was *múltaque*.

Cicero's increasing preference for the type in which *-que* is added to a disyllabic word of form ⏑⏑ is largely confined to the form *eaque*, occurring seventeen times in Shipley's corpus. It is noticeable that, of the other words of this type which occur in the speeches of Cicero examined by Shipley, four more consist of pronominal adjectives plus *-que* (two examples of *suaque* and two of *meaque*, all unelided). In one further example *-que* is added to a preposition (*sineque*, with elision).[10] Five examples are of unelided *utique* 'and that', all of them occurring in decrees and evidently belonging to the language of decrees. The remaining examples are of *uiaque* (once, unelided) and *bonaque* (once, with elision). It is striking that the vast majority of examples of the type concern the addition of *-que* to a pronominal form; it is likely that pronouns were unaccented when unemphatic (see Adams 1996), and therefore that, like prepositions, they will have remained unaccented when followed by *-que* (see section 2 above). The same applies, of course, to the preposition *sine* followed by *-que*. In other words, these are forms to which the grammarians' rule would not have applied. Examples of the type which should have undergone the grammarians' rule are therefore very few, confined in Shipley's corpus to *utique* 'and that' in the language of decrees (but on the length of the *-i-* in *utique*, see section 2 above) and otherwise to one example of unelided *uiaque* and one of elided *bonaque*.[11]

It appears, then, that Cicero avoided adding *-que* to words ending in a short vowel, except, in his later speeches, for certain sequences of the form ⏑⏑ + *-que*. Sequences of the latter type which occur were mostly unaccented.[12] If the normal Latin rule for accenting single full words applied in Cicero's time to combinations of full word plus enclitic, and if Cicero's avoidance of short vowel + *-que* sequences is related to their accentuation, the avoidance can be explained only as the result of various different factors applying to different word shapes.

---

[10] The occurrence is at *Pro Caecina* 64, which is not in Shipley's corpus, but Shipley (1913: 39), apparently misled by an abbreviation such as 'Caec.', mistakenly gives the passage as *Divinatio in Caecilium* 64. I mention the form since it is discussed by Shipley, though its occurrence ought to be deleted from his statistics.

[11] The accentual effect of an elided enclitic is a question which cannot be considered here. But cf. passage (11) above.

[12] Shipley leaves out of consideration *ítaque* 'therefore' and *itáque* 'and thus'. Instances of *ítaque* 'therefore', which occur frequently, were of course, according to the grammarians, accented but not subject to the grammarians' rule.

Apart from the circularity we have noted in Shipley's attempt to elucidate these factors, it obliges us to assume that Cicero's general avoidance of the sequence short vowel + *-que* is coincidental: the reason for avoidance of the type *multaque* is quite different from the reason for avoidance of the types *crīminaque* and *positaque*. If, on the other hand, the grammarians' rule applied in Cicero's time, there would at least be some unity to what was avoided: all sequences of accented short vowel followed immediately by *-que*, that is to say all sequences of short vowel plus *-que* except (at least in the later speeches) for certain sequences of the form ⏑⏑ + *-que* which were unaccented (or, in the case of *ítaque* 'therefore', accented but not on the syllable immediately preceding *-que*: see n. 12). It is less clear *why* sequences of accented short vowel + *-que* should have been avoided. Possibly, the solution is to be sought not in authors' rhythmic preferences but in the general type of situation which can give rise to avoidance of a linguistic form. If so, one might speculate that the grammarians' rule was in the process of replacing some earlier rule for accenting such words, in which case a period of uncertainty about their accentuation may have manifested itself in a tendency to avoid such sequences.[13] As long as we do not have direct evidence for such a development, however, this must remain speculative. What is somewhat clearer is that a unified account of Cicero's avoidance of *-que* in certain combinations is much more likely to emerge under the assumption that the grammarians' rule was operative, whether in general or as a relatively new possibility.

5. ítaque *versus* itáque *revisited*

We have seen that the grammarians are remarkably consistent in their descriptions of the accentual effects of enclitics. The enclitics they recognise are *-que*, *-ve*, *-ne*, and *-ce*. The main difficulty confronting them is that certain words in which *-que* does not mean 'and' are accented according to the enclitic rule (*utráque, plēráque*), whereas others are not (*ítaque* 'therefore', *úndique, dénique*).[14]

Historically, we know that the *-que* of all these words has the same source as the *-que* meaning 'and'. If, therefore, the grammarians' rule for the accentuation of enclitics is correct, we might expect that all words ending in *-que*, with

---

[13] For the avoidance of forms as a stage in their replacement, see Pinker (1999: 125–6) = Pinker (2000: 139–40), Dickey (2000: esp. 554–5, with n. 31).

[14] I leave out of account here *útique* 'certainly', *utíque* 'and that', and *ubíque* 'everywhere', because of the uncertainties as to the quantity of the *-i-* in the penultimate syllable: see section 2 above.

whatever meaning, should be accented on the syllable preceding *-que*: *utráque*, *plēráque*, **itáque* 'therefore' (as well as *itáque* 'and thus'), **undíque*, **dēníque*. If some, but not all, of the words in which *-que* does not mean 'and' are accented instead according to the normal Latin rule for the accentuation of full words, this suggests that some combinations in which *-que* had lost the meaning 'and' had ceased to be treated as combinations of full word plus enclitic and had come to be treated instead as full words in their own right. It is easy to see how this might have happened. The element *-que* is a structural unit which is normally identifiable on the basis of its phonological form (*-que*) combined with its meaning ('and'). In cases where the meaning 'and' has been lost, the element has lost one of its identifying characteristics and may cease to be analysed as a separate element of the word: once a sequence such as *itaque* is no longer obviously composed, as far as its meaning is concerned, from *ita* 'thus' and *-que* 'and', the way is open for reanalysis as simply a full word, with no boundary between *ita* and *-que*. But once so reanalysed, a form such as **itáque* 'therefore' would have been highly irregular in its accentuation, since full words in Latin are, with very few exceptions, accented according to the rule that the accent falls on the penultimate syllable, if heavy, and otherwise on the antepenultimate.[15] It is therefore entirely unsurprising that the accentuation of such a form should have been regularized, giving *ítaque* 'therefore'.[16]

Instances in which accentuation is affected by a change in the synchronic analysis of a particular sequence can be adduced from other languages. In English, a phrase such as *blàck bírd* may be differently accented from a compound such as *bláckbird*. In the case where the meaning cannot wholly be predicted from the meanings of the component parts, the sequence has come to be treated not as a phrase in which *black* is an adjective modifying *bird* but simply as a word, and its accent follows the most regularly applicable pattern for

---

[15] There is a handful of exceptions which arose historically when the final syllable was lost and the syllable which had been the penultimate, and had carried the accent, became final but did not lose its accent, e.g. *illĭc* 'there'< **illĭce*: see Niedermann (1953: 14).

[16] Richard Ashdowne points out to me that a stage with **itáque* 'therefore' did not necessarily exist as such, especially as early Latin had a different accent system altogether, with accentuation fixed on the initial syllable of a word. We do not know how the accentuation of a word was affected at this period by a following enclitic, but it is altogether possible that the accentual effect of enclitics described by the grammarians arose no earlier than the change to the Classical Latin rule for assigning accents. If so, reanalysis of a sequence such as *itaque* 'therefore' as consisting simply of a full word may have arisen during the period of the early Latin initial accent, and *itaque* may have emerged at the change to the Classical Latin accent rule as a fully-fledged ordinary word with accent simply assigned to the antepenultimate syllable according to the normal rule that was coming into force. I do not think the main argument is affected if **itáque* 'therefore' did not exist as such, but the actual sequence of events would be slightly different if it did not.

disyllabic nouns in English, i.e. accentuation on the first syllable. The internal morphological structure of a single word can also affect its accentuation in English. For example, certain derivational suffixes are associated with an accent in a certain position. Historically, many English words in *-ate* have undergone a change from accentuation on the penultimate syllable to final accentuation if disyllabic or antepenultimate accentuation (with a secondary accent on the final syllable) if trisyllabic: díctate => dictáte; concéntrate => cóncentràte (see Gasiorowski 1997). Phillips (1998*a*, 1998*b*) has argued plausibly that the stress shift occurred first in those words in which *-ate* was least likely to be synchronically analysed as a suffix, and that after this loss of synchronic morphological analysis a word tended to acquire an accent that was more regular for synchronically unanalysed disyllabic and trisyllabic verbs.[17] Post-archaic Latin, with its very straightforward rule for assigning an accent to a word, does not have many opportunities for change in the synchronic analysis of a sequence to have consequences for its accentuation. The presence or absence of a boundary between full word and enclitic does, however, affect the accentuation of a Latin word, if we trust the grammarians' rule. It is here that, if the rule is valid, we would expect to see accentual effects of the loss of the synchronic sensation of a boundary, and the minimal pair *ítaque* 'therefore' versus *itáque* 'and thus' is therefore entirely to be expected.

The significance of the minimal pair *ítaque* 'therefore' versus *itáque* 'and thus' is that, although the difference in accent between these forms is entirely to be expected, the grammarians were unable to explain it straightforwardly. They could suggest that *-que* was not an enclitic when it did not mean 'and', but this caused difficulties with *plēráque* and *utráque*. One could resort to the suggestion that *plēráque* and *utráque* were accented by analogy with the corresponding masculines and neuters, but this meant explaining *plēráque* and *utráque* quite differently from other sequences in which *-que* was immediately preceded by an accented syllable (*doctúsque*, *līmináque*, *itáque* 'and thus'). In addition, as far as our evidence allows us to tell, paradigmatically related forms did not normally affect each other's accentuation in Latin. Doubt lingered as to the explanation for *plēráque* and *utráque*: perhaps the accentuation of these sequences had, after all, something to do with the accentual effect of an enclitic. But then, how could one possibly define the circumstances under which *-que* was an enclitic and those under which it was not? Priscian in passage (9) hints at some highly credible possibilities: that it makes a difference whether *-que* is added to a base that can be recognized synchronically as a word in its own right, and that it makes a difference whether only the element *-que* has lost its proper significance

---

[17] I have argued elsewhere for similar phenomena in ancient Greek (Probert 2000).

(as in *utráque*) or whether the base to which *-que* is attached has also been deprived of its basic meaning (as in *ítaque* 'therefore'). Priscian displays here an instinctive understanding that the more formally and semantically transparent a combination of full word plus *-que*, the more likely it is that the accent will fall on the syllable immediately preceding *-que*. The fact that *-que* is added to a base that can be recognized synchronically as a whole word contributes to transparency, but the fact that a base has lost its proper meaning detracts from transparency. Thus, *plēráque* is not totally transparent, since *-que* has lost its proper meaning, but is more transparent than *ítaque* 'therefore', in which both base and *-que* have lost their proper meanings. An even less transparent sequence is *dénique*, in which the base cannot even be recognized synchronically as a whole word. The point at which a sequence ceases to be analysed as a sequence of full word plus *-que* appears to be the point reached with *ítaque* 'therefore'. A sequence at least as opaque as *ítaque* 'therefore' is simply accented as a full word; a more transparent sequence is accented according to the grammarians' rule for enclitic accentuation. This connection between the transparency of a sequence and its accentuation as the relevant type of sequence is excellently supported typologically, but the Latin grammarians do not have the typological knowledge to see the credibility of such a connection. Instead, they are puzzled by the messy data at their disposal, they grope around for explanations, and they are unable to settle definitely for any one explanation. The typological plausibility of the data we are presented with, combined with the fact that the grammarians do not have the knowledge that would have led them to invent precisely these data, contributes immensely to the credibility of the data themselves, and with the data to the fundamental principle of enclitic accentuation described by the grammarians.

To sum up, the minimal pair *ítaque* 'therefore' versus *itáque* 'and thus', combined with some related data supplied by the Latin grammarians on enclitic accentuation, makes it highly likely that the grammarians' rule for enclitic accentuation is neither invented on the basis of Greek enclitic accentuation nor due to over-generalization from cases such as *doctúsque*, but is a linguistically real description of the facts of Latin. It remains to ask for what periods of Latin we can assume the grammarians' rule for enclitic accentuation to have been in force. Here some caution is necessary. The extant statements of the grammarians' rule are all post-Classical and therefore do not necessarily provide any information as to the accentual effects of enclitics in the Classical period. As mentioned above (see section 2, with n. 5), the ultimate source of Martianus Capella's statement of the enclitic rule (passage (6) above) has been taken to be Varro. If this inference is correct, the enclitic rule was stated at least as early as the first century BC and was therefore valid at the time of Cicero and Vergil. We

have seen some independent support for the notion that the enclitic accent rule was in force at this period in Cicero's avoidance of sequences which under the relevant rule would have consisted of accented short vowel plus *-que* (section 3 above). On the other hand, we have also wondered whether Cicero's avoidance of such sequences points to some uncertainty in his time as to their accentuation. I leave open the question as to the exact situation for the Classical period, but hope to have strengthened the case for regarding the grammarians' rule for enclitic accentuation as linguistically real, and for exploring with some confidence in the validity of the rule the question as to exactly when it was in operation.

References

Adams, James Noel (1996). 'Interpuncts as evidence for the enclitic character of personal pronouns in Latin'. *Zeitschrift für Papyrologie und Epigraphik*, 111: 208–10.

Allen, William Sidney (1989). *Vox Latina* (2nd edn., corrected reprint). Cambridge.

Corssen, Wilhelm Paul (1858–9). *Über Aussprache, Vokalismus und Betonung der lateinischen Sprache* (2 vols.). Leipzig.

Dickey, Eleanor (2000). '*O egregie grammatice*: the vocative problems of Latin words ending in *-ius*'. *Classical Quarterly*, 50: 548–62.

Gasiorowski, Piotr (1997). 'Words in *-ate* and the history of English stress', in Jacek Fisiak (ed.), *Studies in Middle English Linguistics*. Berlin, 157–80.

Halle, Morris (1997). 'On stress and accent in Indo-European'. *Language*, 73: 275–313.

Kent, Roland Grubb (1945). *The Sounds of Latin* (3rd edn.). Baltimore, Md.

Liénard, Edmond (1969). 'Réflexions sur l'accent latin', in Jacqueline Bibauw (ed.), *Hommages à Marcel Renard*. Bruxelles, 551–60.

Niedermann, Max (1953). *Précis de phonétique historique du Latin* (4th edn.). Paris.

Phillips, Betty (1998*a*). 'Word frequency and lexical diffusion in English stress shifts', in Richard Hogg and L. van Bergen (eds.), *Proceedings of the Twelfth International Conference on Historical Linguistics*. Philadelphia, 223–32.

—— (1998*b*). 'Lexical diffusion is *not* lexical analogy'. *Word*, 49: 369–81.

Pinker, Steven (1999). *Words and Rules: The Ingredients of Language*. London: Weidenfeld and Nicholson.

—— (2000). *Words and Rules: The Ingredients of Language*. London: Phoenix.

Probert, Philomen (2000). 'Studies in ancient Greek accentuation'. DPhil Dissertation, Oxford.

Schöll, Fridericus (1876). *De accentu linguae latinae veterum grammaticorum testimonia*. Leipzig.

Shipley, Frederick William (1909). 'The effect of enclitics on Latin word accent in the light of Republican prose usage'. *Proceedings of the American Philological Association*, 40: lxxxiii–lxxxiv.

—— (1911*a*). 'The treatment of dactylic words in the rhythmic prose of Cicero'. *Transactions of the American Philological Association*, 41: 139–56.

—— (1911*b*). 'The heroic clausula in Cicero and Quintilian'. *Classical Philology*, 6: 410–18.

—— (1913). 'Preferred and avoided combinations of the enclitic *-que* in Cicero'. *Classical Philology*, 8: 23–47.

Soubiran, Jean (1966). *L'élision dans la poésie latine*. Paris.

Steriade, Donca (1988). 'Greek accent: a case for preserving structure'. *Linguistic Inquiry*, 19: 271–314.

Tucker, Robert Whitney (1965). 'Accentuation before enclitics in Latin'. *Transactions of the American Philological Association*, 96: 449–61.

Wagener, Carl (1904). 'Betonung der mit *que*, *ve*, *ne* zusammengesetzten Wörter im Lateinischen'. *Neue philologische Rundschau*, 505–11.

The linguistic basis of textual segmentation
in Serbian Church Slavonic sources of the 14th-15th centuries

C. M. MacRobert

Early Cyrillic manuscripts are frustratingly uninformative about the prosodic organization of the texts which they contain. For the first three hundred years from which they survive (11th-13th c.) they lack stress-marks. These came gradually and at first sporadically into use from the late 13th c. onwards (Nedeljković 1967: 23-8).[1] For the most part scribes follow the Greek practice of *scriptio continua*: points are used to divide the text into periods or paragraphs, clauses or phrases, but the layout does not acknowledge subdivisions within these units, though breathings may be used to mark vowels in syllable-initial position. It is commonly asserted in works on Cyrillic palaeography (e.g. Karskij 1928: 236) that this state of affairs held approximately until the 16th c., when subdivision into prosodic units — stress-groups — came into use; only Đorđić (1971: 183) and Grickat (1961-2: 244) suggest a rather earlier starting point for this practice in Serbian and especially Bosnian manuscripts, from the 14th c. onwards, and Đorđic offers a summary account of the basis for this type of segmentation: prepositions, conjunctions and all proclitics and enclitics in general are written together with the words (i.e. stressed words) next to which they occur.

These statements call for some qualification. In the first place, there is a small amount of early evidence for segmentation of text into word-size units, probably on a prosodic basis, from the western edge of the Cyrillic/Glagolitic area. The 11th-c. Freising Fragments (Kos 1993), three early Slavonic liturgical texts written in Latin letters, reflect the Western European practice of word-separation which had long since spread from Insular to Continental use (Parkes 1992: 23-6, Saenger 1997: 30-2).[2] Of more significance is the segmentation into word-size units in the Kiev Folia (Trubetzkoy 1954/1968: 58, Schaeken 1987: 21-2), a fragmentary Glagolitic manuscript associated on linguistic grounds with Moravia and conventionally attributed to the 11th c. The other early Glagolitic

---

[1] Some information about stress-patterns has been inferred from early liturgical manuscripts which include musical notation by Störmer (1987).

[2] The fact that the segmentation is sometimes inept suggests that the scribes were trained in a different linguistic tradition, probably German/Latin, and had an imperfect understanding of the Slavonic text which they wrote.

manuscripts, which are thought to be of Bulgarian or Macedonian provenance, are apparently written in *scriptio continua*,[3] but the Glagolitic writing tradition which established itself further west, in Istria, Dalmatia and their off-shore islands, observed segmentation of text into prosodic units at least from the 13th c., as can be seen from the fragmentary manuscripts extant from that time (Vajs 1932: plates xxviii-xxx).

Secondly, there is a growing body of evidence for segmentation of text below the level of the clause or phrase in Cyrillic sources before the 16th c. Part of this evidence comes from the northernmost part of the area in which Cyrillic was used, from the Novgorod birchbark texts which have come to light in archaeological excavations from the 1950s onwards and are still being unearthed year by year. The use of points in these short and relatively informal documents is so variable that it is not readily amenable to analysis, though Zaliznyak (1995) notes some of the more unusual distributions in his commentaries. Those birchbark texts which are long enough to contain a number of points seem to fall into five reasonably distinct types:

(i)   those written in pure *scriptio continua*;[4]
(ii)   those written in *scriptio continua*, but with points used to delimit clauses (and sometimes phrases);[5]
(iii) those which consist of a series of parallel items, typically nouns or phrases, separated from each other by means of points or other markers;[6]
(iv) those in which single or double points are apparently used to delimit prosodic units (what Zaliznyak calls 'phonetic words');[7]
(v)   those in which points, especially double ones, are used to separate syllables.[8]

The use of points in the Novgorod birchbarks to mark prosodic units seems to be predominantly a feature of vernacular texts such as private letters and wills, and

---

[3] Sporadic segmentation, probably on a prosodic basis, is indicated by spacing in the single-column portion of the Codex Assemanianus (Ivanova-Mavrodinova and Džurova 1981) and by the use of double points in the Grškovićev fragment (Trifunović 2001: 26 and plate).

[4] E.g. nos. 605, 682, 657, 717 (12th c. monastic letters), 531, 705 (12th-13th c. family letters), 69 (late 13th-c. business letter), 497, 19 (14th-15th c. family letters).

[5] E.g. nos. 644, 424, 731 (12th c. family letters), 222, 420 (13th c. business letters), 142, 354, 358, 99, 610, 272 (14th c. business letters), 49 (14th-15th c. family letter), 310 (15th c. business letter).

[6] E.g. nos. 671, 601, 609, 438, 390, 138, 92 (12th-14th c.).

[7] E.g. nos. 109 (11th-12th c. business letter), 335 (12th c. business letter), 615 (13th c. buiness letter), 140, 5, 102 (14th c. business letters), 578 (14th c. family letter), 370, 446, 178, 167, 275/266, 248, 249, 314 (14th c. business letters), 125, 124, 43 (14th-15th c. family letters), 519, 692 (14th-15th c. wills).

[8] Most obviously in nos. 422, 112, 707, 412 (12th-13th c. business letters).

is paralleled in some legal documents on parchment or paper from the same area.[9] However, the same practice is also found in at least one Church Slavonic manuscript of Novgorod provenance, an early 15th-c. psalter.[10]

The other area in which textual segmentation below the level of the clause or phrase can be observed in Cyrillic manuscripts before the 16th c. is also geographically peripheral: in the south-west Balkans the practice of leaving spaces between prosodic units is attested sporadically from the later 13th c. onwards, both in documents and in church books. As Đorđić (1971: 183) observes, it is frequently to be seen in the small corpus of manuscripts which on palaeographical, linguistic or historical grounds are thought to have been written in Bosnia. The dating of these manuscripts, which is in most instances at best approximate (Lavrov 1914: 236-49, Grickat 1961-2: 244, Đorđić 1971: 129-38), ranges from the 13th c.[11] through the 14th c.[12] to the mid 15th c.[13] Only a minority are written in *scriptio continua*.[14] Several of the manuscripts seem to have been written under Glagolitic influence (Kuna 1977), which could be the source from which the practice of textual segmentation was derived.

In general, however, the palaeographical, ornamental, orthographical, and textual peculiarities of the Bosnian Cyrillic manuscripts seem to represent a conservative tradition (linked in its turn to earlier Glagolitic) whose closest parallels are found in a few manuscripts of the 12th-13th c. from the central and south-western Balkan areas, such as the Dobromir (Altbauer 1973), Miroslav (Rodić and Jovanović 1986) and Vukan (Vrana 1967) Gospels. These manuscripts are written in *scriptio continua*; but the practice of textual segmentation into prosodic units was known in the late 13th and early 14th century more widely than Bosnia. It can be seen in some examples of 13th and

---

[9] E.g. the will of Kliment written before 1270 (Obnorskij and Barxudarov 1952: 54-6), the trading treaty between Polotsk and Riga of 1330 (ibid., 93), or the Novgorod document of 1412 (Zaliznjak 1995: 373-8).

[10] No. 35 from the collection of the Typography of the Holy Synod in Moscow, now in the Russian State Archive for Ancient Documents (RGADA).

[11] The Grigorovič/Hilferding Gospel fragments (*pace* Grickat 1961-2: 244; see Mošin 1966: plate 125).

[12] The Manojlovo Jevandelje (Speranskij 1906: 12-17), the Dovoljsko Jevandelje (Mošin 1966: plate 130).

[13] The Batalovo Jevandelje of 1393 (Mošin 1966: plate 132), the Panteleimon Gospel (Tachiaos 1981: 30-2 and plate 2), the Nikolsko Jevandelje (Daničić 1864: plates), the Čajnicko Jevandelje (Jerković 1975: plates), the Kopitarevo Jevandelje (Šidak 1955: plates), the Pripkovićevo Jevandelje (Mošin 1966: plate 127), the Hvalov sbornik of 1404 (Kuna 1986), and Radosav's manuscript (Džurova *et al.* 1985: plates cci-cciv).

[14] The Leningrad Gospel (Mošin 1966: plate 126) and apparently the Divoševo Jevandelje (Grickat 1961-2: plates).

14th-c. Serbian diplomatic.[15] A search through recent catalogues of manuscripts discovers it in some Serbian church books preserved in the monasteries of Mt Athos[16] and in some manuscripts which have found their way to Russia.[17] There are also examples among liturgical books of the early 14th century in Belgrade National Library[18] and in an early 14th-c. psalter, No.68 from the manuscript collection of the Patriarchate in Peć.[19] In the later 14th c. the practice of dividing text into prosodic units seems to become less common, at any rate in church books, probably because of renewed Greek and Bulgarian graphic and orthographic influence, but it lingered in some places[20] and re-emerges in early printed books, such as the Cetinje Psalter published by Đurđe Crnojevicv in 1494.

Unfortunately most of these Bosnian and Serbian sources have not been studied in depth or published in full; photographs of selected folia allow one to see that the scribes and printers divided their text into units which were probably determined by stress-patterns, but they do not provide enough data to allow detailed characterization of those patterns. When manuscripts have been published in transcription, rather than facsimile, modern word separations have been imposed by the editor without regard for the scribe's segmentation of the text. Even the admirably systematic and detailed studies of scribal practice by Jerković (1975), Ivić and Jerković (1981 and 1982), and Grković-Mejdžor (1993) provide little or no information on the subject of prosodic units. The analysis which follows is therefore based only on the material which is available to me *in extenso*: the psalter manuscript Peć 68, which I am preparing for publication from microfilm, and the modern facsimile of the Cetinje Psalter of 1494 (Martinović 1986).

For the study of low-level textual segmentation these two sources have some advantages, but also present certain drawbacks for which allowance has to be made. Their primary advantage is their length: the part of manuscript Peć 68,

---

[15] See Pavlović (1958), Đorđić (1971: illustrations 58-61, 338, illustrations 135-6), and particularly King Milutin's charters of 1302 (Mošin 1966: plate 67) and 1313-16 (Jagić 1890; Đorđić 1971: 283, illustration 69).

[16] The liturgical roll Hilandar 16/5 of the late 13th century (Bogdanović 1978: plate 18), the 1316 tetraevangelion of King Milutin, Hilandar 1 (Bogdanović 1978: plate 25), Roman's colophon to his Gospel Lectionary of 1337, Hilandar 9 (Bogdanović 1978: plate 28), a Gospel lectionary, Panteleimon 2, of the 13th-14th century (Tachiaos 1981: plate 1).

[17] The 14th-c. Athonite Gospel book F.I.87 (Mošin 1966: plate 75) and the collection of Lenten discourses of S. Isaac of Syria, Pogod. 72 (Mošin 1966: plate 92), both in the Russian National Library in St Petersburg.

[18] Nos. 3, 93, 487, 643-4 (Štavljanin-Đorđević *et al.* 1991: plates 30, 39-40, 43, 48).

[19] See MacRobert (1989: plates).

[20] For an example from 1408, see Mošin (1966: plate 137).

folia 1r-185v, written by the main scribe, who was in the habit of indicating prosodic units, comprises psalm 33:7 to the end of psalm 149 plus the first three Old Testament Canticles[21] and a number of hymns and prayers inserted at specific intervals. This body of text, when subjected to modern word-separation, amounts to a corpus of approximately 29,200 words. The Cetinje Psalter is a much longer book, which incorporates a large amount of liturgical material, but for comparative purposes I have used only the psalms and canticles which correspond to those listed above from Peć 68. The other advantages of these two sources are their clear and careful production, which assists interpretation, and the character of the text which they contain. The Church Slavonic Psalter is subject to a traditional division into verses and versicles which is taken from the Greek Septuagint and varies little from one manuscript to another. It was one of the most widely read and recited ecclesiastical texts, used as a primer for the acquisition of literacy, known to clerics by heart. Whether or not scribes fully understood the text of the Psalms, they were likely to be thoroughly familiar with it and to agree about how to divide it into semantico-syntactic units. Thus on the whole higher-level textual segmentation is the same in Peć 68 and the Cetinje Psalter, and does not complicate or obscure the examination of lower-level segmentation into prosodic units.

The difficulties which the two sources present stem from the rules governing their layout. Although both the main scribe of Peć 68 and Makarije, the printer of the Cetinje Psalter, frequently used spacing to mark prosodic units, this was not the first priority in the visual organization of text for either of them, and it probably did not have the status of an obligatory orthographical practice. Rather, it belonged to the class of optional rules which the medieval scribe might implement systematically, but whose neglect apparently did not constitute a serious infringement of the norms of literacy (Birnbaum 1988: 129-31, Osipov 1992: 182-3). Essentially, both scribes were guided by traditional rules for the layout of manuscripts intended for liturgical use; the systematic differences between them derive partly from two distinct ways of presenting the psalter text, and partly from the divergent requirements and constraints of the handwritten and the printed book.

The scribe of Peć 68 seems to have observed the following four rules:

(1) The manuscript line must end in a vocalic letter. This rule, which is very widely applied by scribes working in Cyrillic, probably goes back to the introduction of the Cyrillic alphabet in the late 9th-10th c., when Old Church Slavonic was an open-syllable language. In order to satisfy the rule,

---

[21] The Song of Miriam, Exod. 15:1-19, the Song of Moses, Deut. 32:1-43, and the Prayer of Hannah, I Sam. 2:1-10.

the scribe resorts to ligatured and superscript consonantal letters towards the end of the line, employs conventional final vocalic letters which had no phonetic value, divides versicles, prosodic units and words between two lines, and compresses lettering at the expense of the spaces which he elsewhere leaves between prosodic units.

(2) Subject to rule (1), each versicle of the psalms must start on a new line and if possible be contained within that line. To this end, the scribe again uses abbreviatory devices such as ligatures and superscripts; if the versicle is just a little too long to fit easily within the ruled margins which frame the text, he may resort to compression of lettering or allow the line to protrude into the right-hand margin, or both.

(3) Subject to rules (1) and (2), the manuscript line must be confined within the ruled margins. The application of this rule is most apparent in the hymns and prayers which accompany the psalms: the syntactic units (clauses, for the most part) of these texts are marked with points but do not start each on a new line, and are frequently divided between lines in order to preserve the margin. Again, abbreviation and compression can help to achieve this.

(4) Subject to rules (1), (2), and (3), prosodic units are separated from each other by spacing.

The printer Makarije ordered his priorities differently:

(1) The printed line must begin and end at the margins, i.e. it must be justified. This requirement leads to noticeable variations in spacing. Sometimes, especially after the large ornamental initials which take up about a quarter of several lines in succession, letters may be compressed so that space between putative prosodic units is lost; on other occasions, in order to stretch the line out to the margin, as much space may be left between the letters within a word as is elsewhere used to separate prosodic units. The application of this rule is facilitated by printing the versicles of the psalms continuously (though marked off from each other by rubricated points) and allowing them to run over freely from one line to another. Prosodic units, words and even syllables and consonant clusters may also be divided between lines for the sake of justification. This rule is not the same as rule (3) in the manuscript, for the scribe of Peć 68 clearly felt no compulsion to extend the shorter psalm versicles as far as the right-hand margin of his page, and, when he was obliged to divide a versicle between two lines, did not start the second line at the left-hand margin, but indented it.

(2) Subject to rule (1), the printed line must end in a vocalic letter. This is Makarije's predominant practice, and is assisted by the division of textual units (versicles, clauses) and prosodic units between lines, by compression of lettering towards the end of the line. The fact that he occasionally abandons it may simply reflect the difficulty of producing justified text with rather large, hand-set type; but it is also possible that he was working in a local tradition which allowed lines of text to end in consonants, as in the Bosnian and early Serbian and Macedonian manuscripts mentioned above.

(3) Subject to rules (1) and (2), prosodic units are separated from each other by spacing.

It can be seen from these formulations that the recognition of prosodic units in these two sources is to some extent a matter of judgement. At best the size of the spaces which separate one unit from another is not standard, and sometimes it has to be reassessed from one line to another, with allowance made for the number of letters that the scribe or printer has elected to accommodate in each line. In certain places — towards the end of the written or printed line in both sources, towards the end of the verse in Peć 68, after ornate initial letters in the Cetinje Psalter — the letters may be placed so close together as to leave the segmentation indeterminate. The modern transcriber then steers a precarious course between the temptation to read into the smallest gap or variation in the alignment of letters evidence of deliberate spacing (reinforced, as one proceeds, by the growing sense that one 'knows' what constitutes a prosodic unit) and the seductively easy option of taking a continuous string of letters at face value. Neither danger need be pernicious, provided each is recognized. Small differences between the ways in which groups of letters are placed on the line can reflect the junctures at which the scribe paused briefly to consult his examplar, and in the case of a text which he would have known more or less by heart those junctures are quite likely to have coincided with the ends of prosodic units. Unusual patterns of segmentation which result from line-end effects can be allowed for in analysis, provided they are identifiable within the corpus of data.

Fortunately, in the case of Peć 68, the scribe's orthography provides a check on spatial segmentation. He followed a practice already established in the 13th c. (Ivić and Jerković 1981: 73-5, 1982: 79-80) of using the letter *jer* as a final demarcating sign for prosodic units which end in consonants. This is another relic of the open syllable-structure which obtained when the Cyrillic writing system came into being: the letter *jer* formerly stood for a short vowel (originally two vowels, front and back) which was subsequently lost in final

position and a range of medial positions, but continued to be used in writing. Because its distribution was determined by tradition rather than pronunciation, it was liable to be replaced by the apostrophe (*pajerak*) or even to be omitted entirely, especially from consonant clusters at the juncture of stem and suffix. The scribe of Peć 68 treats enclitics as a type of suffix, marking the end of the preceding orthotonic item indifferently with *jer* or with the apostrophe; but at the end of the whole prosodic unit he admits only *jer*, except at the end of the written line, where the apostrophe may be used (x 95) to save space. There is only one example in the whole of his work where a prosodic unit ends in an apostrophe within the line: *Blg{s*}n' g*ĭ+b*ĭ+moi* (166r, 15).[22] It is even more unusual for a prosodic unit to end in a consonant without a following *jer* or apostrophe, and the two clear examples are probably motivated by considerations of spacing, since they occur at the ends of lines. Thus a word-form whose final letter is a consonant or apostrophe can safely be regarded as a non-final element in a prosodic unit. This correlation can help to resolve some indeterminacies in spacing, though not, of course, the frequent cases where a word-form ends in a vocalic letter (including *jer*).

A further difficulty in Peć 68 is that the text which the scribe wrote is sporadically obscured by the interventions of a later corrector, who went through the whole psalter erasing forms, words, and phrases and substituting a 14th-c. revised version. Although the original wording can often be made out or inferred from other early manuscripts on the basis of the space which it occupied, in some instances the segmentation is left in doubt. Fortunately again, the corrector confined his attention to the psalms, making no interventions in the accompanying hymns and prayers. These texts, which make up about 3,840 words out of the total 29,200 mentioned above when modern separation is applied, offer a clearer picture of scribal segmentation not only because they are free of later corrections, but also because they are subject only to the scribe's rules (1), (3), and (4) outlined above. If allowance is made for the effect of rule (4), they turn out to contain about 2,650 prosodic units, of which a good third, about 940, are composite, consisting of more than one item which in modern terms would be regarded as a separate word. The supplementary texts thus provide a data-base which is modest in size and scope — 1st and 2nd person singular forms predominate, there is little use of past tense verbal forms — but

---

[22] Scribal diacritics are omitted from the transliterated examples. The *titlo* which marks abbreviated *nomina sacra* is represented by the asterisk; superscript letters are enclosed in braces; a plus sign is used to link lexical items which would be separated by spacing in modern practice, but appear without separation in the sources; a single vertical line is used to indicate a space within a compound word; a double vertical line indicates line-end in the manuscript. References are supplied by folio and line.

sufficient to allow preliminary conclusions, which can then be tested against the larger and more diverse body of data contained in the text of the psalms.

Four types of prosodic unit emerge clearly from the supplementary texts. Firstly, there are the items which stand alone, being apparently orthotonic. They make up the greater part of the vocabulary, but many of them occur only once. Secondly, there are the items which frequently and consistently co-occur with orthotonic items in spatially demarcated prosodic units and so can be assumed to have been clitics. They can be divided into two main groups, each of which can be subdivided into a series of sequentially ordered categories:

(1) proclitics — the co-ordinating conjunctions *a*, *i*, *ili*, *nĭ*, the subordinating conjunction *da*, the negative particles *ne* and *ni*, the monosyllabic and asyllabic prepositions *do*, *kĭ/k*, *iz*, *na*, *ot*, *po*, *pod*, *prě*, *prěd*, *pri*, *sĭ/s*, *vĭ/v*, *za* (the preposition *bez* is problematic, since it occurs twice as part of a prosodic unit and twice free-standing);

(2) enclitics, i.e. items which do not occur at the beginning of a prosodic unit, but are found only after proclitics or orthotonic items — the clausal particles *bo*, *li*, *že*, the 1st and 2nd person singular and 1st person plural accusative pronouns *me*, *te*, *ny*, the reflexive accusative pronoun *se*, the 1st and 2nd person dative pronouns *mi*, *ti*.[23]

Naturally, the weight of evidence for clitic status depends on the frequency with which the various candidates occur in the text, but it is also significant that they follow the same distributional patterns irrespective of frequency. So *i* is treated as a proclitic, i.e. written together with a following word (x 320), and the small number of instances (x 11) which break this rule all occur at line-end or in headings. The same applies, on a smaller scale, to *nĭ* (proclitic x 26, separated by line-end x 3), *da* (proclitic x 37, separated by line-end x 6), *vĭ/v'* (proclitic x 78, separated by line-end x 2), *na* (proclitic x 49, separated by line-end x 1), *sĭ/s* (proclitic x 14, separated by line-end x 1) etc. The consistency with which enclitics are treated is equally compelling: *se* is written as a enclitic together with a preceding word x 80, with a single instance where it stands alone because it appears at the beginning of a line and so is separated from the orthotonic item at the end of the preceding line; *me* (enclitic x 74, separated by line-end x 5), *mi*

---

[23] In practice the enclitic pronominal forms are normally postposed to the governing verb, e.g. *i+da+da{s*}+mi+se* (94v, 6). This is the usual order in Old Church Slavonic (Večerka 1989: 47-51). The other option is to place the enclitic pronoun after the first stressed item in the clause or phrase, but this occurs only in a few instances, mostly where a dative form is used as a possessive (Večerka 1989: 56), e.g.: *i+sam'+bo+se prě-||ž{d}e suda ōsudixĭ* (56r, 8-9)*, Tebě+se m*limĭ* (173r, 14), *b*e ježe+ti sĭgrě-||šixĭ* (56r, 13-14); *Pomošt'nikĭ+mi budi* (93r, 7), *myslĭ+mi||podaje* (122v, 6-7)*, ōka-||n'nu+mi d*šu* (123r, 5-6), *protiv'nago+mi suposta-||ta* (123v, 14-15), *vĭ+poslěd'njuju+mi pogybě-||lĭ* (173r, 4-5)*.

(enclitic x 26, separated by line-end x 4), *te* (enclitic x 14, separated by line-end x 2), *ti* (enclitic x 24, separated by line-end x 2) behave similarly, while *bo* (x 27) and *že* (x 18) are always written together with the preceding word. The exceptions to the rule which cannot be explained by considerations of lay-out on the page are minimal: *ny* is written together x 5, in isolation x 1: *kaže ny na+pokajanije* (175v, 5); the negative particle *ne* (proclitic x 59, separated by line-end x 2) is occasionally followed by a space when it is the last in a sequence of clitics: *da/|ne* (21v, 4), *nĭ+ne* (175v, 3), *i+da+ne* (93r, 12; 124r, 15).

Thirdly, there is a small range of items which from the modern point of view would count as compound words, but which in the scribe's practice are frequently split between two prosodic units, presumably because they were pronounced with two stresses (possibly one primary and one secondary). This practice is attested for 21 lexical items (*bezĭ/ōtvětĭnĭ* x 1, *bl*go/věstvovanije* x 1, *bl*go/dariti* x 1, *bl*go/srĭdije* x 1, *bl*go/uxanije* x 3, *bl*go/ōbrazno* x 1, *b*go/ukrašennaja* x 1, *člv*ko/ljubĭcĭ* x 8, *člv*ko/ljubije* x 3, *četvrě/dnevĭnĭ* x 1, *dlĭgo/trĭpije* x 2, *dlĭgo/trĭpiti* x 2, *dlĭgo/trĭpělivĭ* x 2, *jedino/čedĭ* x 1, *vĭně/judu* x 1, *vse/čestĭno* x 1, *vse/dětelnĭ* x 1, *vse/moguštĭ* x 1, *vse/moštnyi* x 1, *vysoko/uglĭnĭ* x 1, *xvalo/slovije* x 1), making a total of 35 clear instances, plus five places where the division within the compound coincides with the end of the written line (*bl*go//darešt a*, *člv*ko//ljubĭcĭ*, *člv*ko//ljubije* x 2, *dlĭg o / /t rĭpělive*). However, there is some inconsistency in the treatment of the most common expression, *člv*koljubĭcĭ*, which, though split x 8, is also written as a single unit x 6. The derived verb *člv*koljubstvovati* also occurs once as a single unit. Similarly *bl*godariti* is split, but the single examples of *bl*godarovannĭ* and *bl*godarstvovati* are written without internal division, and most of the compounds whose first element is *bez* are not divided, e.g. *bezvodije*, *bez'grěš'nĭ*, *bezĭzatupljenĭ*, *beskonĭč'nĭ*, *beslovesnĭ*, *bečislĭnĭ*. These discrepancies imply that the rule for marking secondary stress within compounds by means of spacing was optional. There are some other compound words which might seem to be candidates for internal division, but are written only as single units: *bl*govoliti*, *dobrodětelĭ*, *ljuboděistvo*, *ljubodějanije*, *mnogom*lstivĭ* x 2, *mytojem'čĭ*, *slastostrastije*, *žestosrdije*. However, as most of these occur only once in the corpus under consideration, their evidential value is limited.

Finally, there are lexical items which are treated variably: sometimes written together with what precedes or follows, sometimes given separate, *ex hypothesi* orthotonic standing. Among the most frequent are the correlative adverb *jako*, which is written together x 30, separately x 22 (x 7 at line-end), and its interrogative counterpart *kako*, written together x 3, separately x 1; the possessive adjectives *moi*, written together x 23, separately x 64, *tvoi*, written together x 13, separately x 107, *svoi*, written together x 3, separately x 16. The

quantifying pronoun *vĭsĭ*, the oblique case personal pronouns *mene*, *tebe*, *m'ně*, *tebě*, *sebe*, the relative pronoun *iže* and the present tense forms of the verb *byti* are predominantly written separately (or together with the clitics discussed above) but occasionally are combined, typically with the verb which they complement or the noun which they qualify, e.g.:

> *i+ne+ōstavi+mene* (23r, 16), *ne+ōstavi mene* (125v, 17; 124r, 12-13), *ne+prězri+mene* (176v, 14), *ne+prězri mene* (125v, 12, 16);
> *i+m*lju+se+tebě* (175v, 10), *m*lju+se tebě* (175r, 14);
> *vsa+isplĭnjaje* (147v, 3), *vsa+ispovědaju* (176v, 13), *vsa+postradati i+vsa strĭpěti* (95r, 14), *v'se+d*ni života//mojego* (177r, 12-13), *vĭ+v'se d*ni//života mojego* (56r, 12), *vĭ+v'se d*ni+života mo//jego* (178v, 1-2).

In addition to these combinations, there are others in which noun and adjective, or noun and dependent noun, or two nouns in apposition, or noun and verb, are written together. Some of them are attested once only, many involve words which occur with high frequency in the corpus, such as *b*ĭ*, *d*nĭ*, *d*ša*, *g*ĭ*, *grěxĭ*, *ime*, *m*ltva*, *rabĭ*, *s*tyi*, *věkĭ*, *vld*ka*, *vrěme*; but none of these items are treated with complete consistency, and there is a fair number of examples where one and the same combination is written together in one place, separated elsewhere.

One might suppose that these inconsistencies reflected the random application of an optional use of spacing — that the scribe aimed on the whole to write stressable lexical items separately, but sometimes neglected to do so — if it were not for three countervailing considerations. Firstly, line-end conventions, which might be expected to override the rules for spacing, seem to be at most a contributory factor, certainly not a determining one, for combinations of items written without separation. Secondly, the combinations which the scribe wrote without separative spacing are usually syntactic units, as can be seen in the example above; if he had combined words together randomly, one would expect to find more combinations which cut across phrases. Thirdly, random lapses into *scriptio continua* ought to affect all lexical items in the same way; but this is not the case, as the figures cited above show: *jako* is more often written together with a following item, whereas possessive adjectives, stressable personal pronouns, the quantifier *vĭsĭ*, the relative pronoun *iže* and the copula/auxiliary *byti* are more frequently separated by spacing.

It thus seems possible that on the whole the scribe meant what he wrote: when he separated two items by spacing, he envisaged that they might each be stressed; when he wrote them together, he expected them to be pronounced as a single prosodic unit. Given the inconsistencies which this fourth type of

prosodic unit exhibits, its motivation will be at best difficult to determine: rhythmic, syntactic, semantic, traditional, or even idiosyncratic factors may all play some part (Zaliznjak 1985: 120-1). It is however worth considering the possibility that some lexical items are at least statistically more likely to enter into prosodic units of this kind because they belong to the category of the so-called *enclinomena* (Birnbaum 1986: 130-1; Dybo 1981: 19; Zaliznjak 1985: 119), i.e. words whose accentual paradigms are made up partly of desinentially stressed forms and partly of forms which are basically unstressed and which therefore combine with other items, clitic or orthotonic, to form single prosodic units (Stang 1978; Dybo 1981: 52). The possible effect of enclinomena can be tentatively illustrated by the following comparison. The noun *slava*, which is traditionally root-stressed and orthotonic in all cases, is written separately (sometimes with clitics) x 9, including acc. sing. x 4, and occurs only once, in the nom. sing., together with the relative pronoun, whose prosodic status itself is debatable: *jemuže+slava po{d}bajetĭ (57r, 14)*. By contrast, the noun *d*ša*, which is traditionally an enclinomenon, unstressed (or stressed only by default) in the acc. and dat. sing. and the nom. and acc. pl., is written separately x 19, including acc. sing. x 7, and together with other possibly stressed items x 6, in acc. sing. x 3, in dat. sing. x 1. The instances are not numerous enough, nor is their treatment consistent enough, to allow firm conclusions, but the fact that in four out of seven instances where *d*ša* enters into this type of prosodic unit it appears in an originally unstressed form suggests that its status as an enclinomenon may still have had some influence on its prosodic treatment.

When these findings, which are supported by a limited textual corpus, are checked against the larger but rather more problematic data-base provided by the text of the psalms in Peć 68, the results are strikingly similar, and the differences serve mostly to clarify points of detail. For instance, there is ample confirmation of the proclitic status of the conjunctions *i* (proclitic x 1833), *da* (proclitic x 158, separated by line-end x 2), *a* (proclitic x 10), and the prepositions *do* (proclitic x 60), *kĭ/k* (proclitic x 74), *na* (proclitic x 380, separated by line-end x 6), *ō* (proclitic x 92, separated by line-end x 1), *ōt* (proclitic x 291, separated by line-end x 3), *po* (proclitic x 93, separated by line-end x 6), *prěd* (proclitic x 69, separated by line-end x 6), *sĭ/s* (proclitic x 78, separated by line-end x 1), *vĭ/v* (proclitic x 906, separated by line-end x 19), as well as the less frequent *iz* (proclitic x 21), *nad* (proclitic x 11), *u* (proclitic x 6), *pod* (proclitic x 7), *pri* (proclitic x 4), *vĭz* (proclitic x 2), *za* (proclitic x 60), *bez* (x 3), *skrozě* (proclitic x 1, separated x 1). Counterexamples, where a gap is left between preposition and noun, are so few that they can probably be attributed to inadvertence:

Bezĭ bezakonija (37v, 17); *do d*še mojeje* (49v, 13); *na puti* (105v, 9); *Naďĭ vsěmi simi* (71r, 10); *prě{d} b*mĭ* (47r, 2), *Prě{d} jefremomĭ* (75v, 11), *prě{d} g*mĭ* (121r, 10), *prě{d} toboju g*i* (145v, 4); *vĭ prit'čju* (50v, 12).

As in the smaller corpus examined above, the negative particles *ne* (proclitic x 370, separated by line-end x 5) and *ni* (proclitic x 44) are occasionally written together with a preceding conjunction but separated from the following item, *i+ne* (x 15), *a+ne* (x 2), *da+ne* (x 5), *jako+ni* (x 1). There is also one instance where the conjunction *nĭ* (otherwise proclitic x 7) and the particle *ne* appear in sequence but separated from each other and from the following expression; here emphasis may motivate the use of spacing:

　　nĭ ne jako||zakonĭ tvoi (139v, 17-140r, 1).

The picture presented by the enclitics is even more clear-cut. The clausal particles are consistently written together with the preceding item: *bo* (enclitic x 24), *li* (enclitic x 15), *že* (enclitic x 123, separated by line-end x 2); so is the postposition *radi* (enclitic x 4), which occurs only once in the corpus of hymns and prayers. The enclitic accusative and dative forms of the personal pronouns, *me* (enclitic x 258, at line-start x 6), *te* (enclitic x 75, at line-start x 1), *s e* (enclitic x 843, at line-start x 15), *i* (enclitic x 71), *je* (enclitic x 112), *ny* (enclitic x 68), *vy* (enclitic x 5), *mi* (enclitic x 40), *ti* (enclitic x 8, at line-start x 1) display a compelling degree of consistency, with only three instances of deviation:

　　vĭprašaxu me (3r, 2), *Sp*si me b*e* (49v, 12); *prizovu te* (164r, 4).

The treatment of compound words tends to confirm the conclusions drawn above, that they can be split in two, presumably because they can be pronounced with two stresses, but that this practice is optional. There are sixteen lexical items which are written at least once with an internal split, giving a total of 30 instances (*bl*go/dějati* x 1, *bl*go/izvoliti* x 1, *bl*go/voliti* x 6, *bl*go/voljenije* x 2, *bl*go/stvoriti* x 1, *bl*go/korěn'nĭ* x 1, *bl*go/stynja* x 1, *desetĭ/strun'nĭ* x 2, *ino/plemen'nikĭ* x 2, *mimo/iti* x 1, *mnogo/plod'nĭ* x 1, *mnogo/prišĭl'stovati* x 1, *rav'no/d*š'ne* x 1, *skoro/pis'cĭ* x 1, *vsa/sĭžigajemaja* x 1, *zakono/prěstup'nikĭ* x 5, *zakono/prěstup'nĭ* x 2). However, some of these are found also found written as single units (*bl*govoliti* x 6, *bl*govoljenije* x 2, *bl*gostynja* x 3, *inoplemen'nikĭ* x 1, *mimoiti* x 3, *vsasĭžigajema* x 1, *zakonoprěstup'nikĭ* x 1). There is also, as before, a range of compound lexical items which might seem to be candidates for internal split but in fact occur as single units (*bl*gověstiti* x 2, *bl*gověstvovati* x 1, *člv*kougod'nikĭ* x 1, *děloměr'nĭ* x 1, *jedinomysljenije* x 1,

jedinomyslĭnĭ x 1, *inočedĭ* x 1, *inorogĭ* x 2, *mimotešti* x 1, *mimoxoditi* x 4, *novoraslĭ* x 1, *ranoprixoditi* x 1, *zaimodav'cĭ* x 1, *zakonoprěstupljenije* x 1, all compounds in *bez-*, *vele-/velĭ -*, and all instances of *bl*gosloviti*).

The larger amount and range of data provided by the psalter text promises evidence about the fourth type of prosodic unit discussed above. Once more the state of affairs found in the supplementary texts is reproduced on a broader scale. The same lexical items are frequently found in this type of combination, though some more can be added to the list, such as *kostĭ*, *ōči*, *putĭ*, *ruka*, *slovo*, *vragĭ*. The same skewed distributions can be seen for very frequent items: *jako* is written separately x 175, together x 350, whereas *tvoj* is written separately x 663, together x 90, *moj* separately x 478, together x 224, *svoj* separately x 240, together x 37, and the present tense forms of the verb *byti*, which are thought originally to have behaved as clitics (Večerka 1989: 51-2) are written separately x 228, together x 69. Scrutiny of line-ends suggests that the constraints which operated there account only for a minority of the combinations written without separative spacing. When the same small test for enclinomena is applied as to the data from the supplementary texts, it turns out that in the text of the psalms the orthotonic noun *slava* is written separately from other potentially stressed items x 36 (including acc. sing. x 15, dat. sing. x 2), and only x 4 together, in the nom. sing. with the demonstrative pronoun *si*, the gen. sing. with the possessive *tvoi*, the acc. sing. and voc. sing. with the verbal forms *vĭzneseĭ* and *vĭstani*.[24] By contrast, the enclinomenon *d*ša* occurs separately x 90, including acc. sing. x 39, dat. sing. x 7, and is written together with other items (not including clitics) x 28, including acc. sing. x 8, dat. sing. x 1, acc. pl. x 1. The difference is just enough to suggest that although *d*ša* is usually treated on a par with orthotonic items, the effects of enclinomenon status still linger.

The Cetinje Psalter clearly follows a tradition in the use of spacing which is similar to that found in Peć 68. This is most apparent in the treatment of compound words: in 18 instances the two sources split the same compound in the same way, and the Cetinje Psalter offers a further 41 examples of split compounds. Some of these do not occur in Peć 68 because they are specific to the textual redaction which the Cetinje Psalter contains: *bl*go/lěpie* (49:2), *dlĭgo/trĭpelivĭ* x 2 (102:8, 144:8), *dlĭgo/lětstvovati* (Odes 2:27), *jedino/rodnyi* (34:17), *kraje/sěkomyi* (113:8), *malo/d*šie* (54:6), *mnogo/m*lostivĭ* x 2 (85:5, 102:8), *meždou/ramija* (67:14), *prěd/iti* x 2 (84:14, 88:15), *vĭspo/plĭznovenije* (55:14), *vĭse/sĭžegaema* x 4, *zakono/polagati* (83:7). Others are common to the two redactions, but are written with internal split either more frequently or

---

[24] The two instances of *slava* written together with verbs are not entirely certain, because they occur at points where the corrector has intervened.

exclusively in the Cetinje Psalter, such as *bl*go/věstiti, člv*ko/ugod'nikĭ, jedino/mysljenije, jedino/myslĭnĭ, ino/plemen'nikĭ, ino/rogĭ, mimo/tešti, novo/rasĺĭ, zaimo/dav'čĭ, zakono/prěstupljenije*, and some of the compounds in *vele-/velĭ* - and in *bl*go-*.

The treatment of prosodic units made up of orthotonic expressions and clitics or possible enclinomena in the Cetinje Psalter is complicated by the printer's overriding concern for the physical lay-out of the page, but bears comparison with what was found in Peć 68. In approximately 4,450 instances the manuscript and the printed book agree in writing items whose clitic status is well established together with (presumably) orthotonic expressions. In about 2,710 places the scribe of Peć 68 writes as a single unit items which the printer Makarije separated with spacing; these account for most of the problematic cases discussed above, involving words such as the possessive adjectives, *jako, vĭsĭ, b*ĭ, g*ĭ, d*ša*, etc., as well as a substantial number of combinations with clitics. There remain 491 instances where Makarije combines items which the scribe of Peć 68 wrote separately. Of these, 76 correspond to line-breaks in the manuscript, and 166 occur in the first six verses of the relevant psalm, and so may be products of the distortion to the printed line caused by the use of large ornamental initial letters; this leaves a mere 249 places unaccounted for. Although Makarije's use of spacing and its relationship to his not entirely systematic placing of stress-marks (Novaković 1877) deserves further investigation, it seems reasonably clear that he continues traditional practice but reflects the beginnings of a shift from the demarcation of prosodic units by means of spacing towards the modern separation of lexical items in writing. Other Serbian Church Slavonic printed books of the late 15th-17th cc. (Medaković 1958) would furnish a substantial source of further material, which invites comparison with the accentological research undertaken by Bulatova (1975, 1977, 1981, 1989) on Serbian Church Slavonic manuscript sources in which stress-marks are used.

It is the manuscript sources of the 13th-14th cc. mentioned above, though, which offer the most intriguing prospects for further investigation, because they allow us a glimpse, even if dim and limited, of prosodic usage in the period before stress-marks became common. The data which have been analysed here show that there was a well defined set of clitics in early 14th-c. Serbian Church Slavonic — and that this set was not identical to the one reconstructed for Old Church Slavonic. They provide good evidence for the existence of double stresses on compound words from an early date. They open up new possibilities for studying the prosodic organization of continuous text and its interplay with traditional, morphologically determined prosodic patterns. To examine in detail all the spatially demarcated combinations in Peć 68 would be a laborious and

difficult undertaking, because one would have to weigh the prosodic status not merely of each lexical item which enters into those combinations, but also the effects likely to have been produced by its appearance in one or other possible form. Some points would probably remain indeterminate, but the results might help to elucidate the interesting though controversial conclusions which Hinrichs (1985, reviewed by Birnbaum 1986) draws about stress-groups in a 14th-c. Bulgarian Church Slavonic psalter manuscript. More generally, the material discussed here demonstrates that in these sources, as in other times and places (Davies 1987), low-level textual segmentation was based on the prosodic unit, in preference to the more abstract lexical items to which modern convention gives pride of place.

References

Altbauer, Moshe (1973). *Dobromirovo evangelie. Kirilski spomenik od XII vek.* Skopje.

Birnbaum, David J. (1986). 'On the methods of analyzing accented Slavic manuscripts'. *International Journal of Slavic Linguistics and Poetics*, 34: 123-43.

Bogdanović, Dimitrije (1978). *Katalog ćirilskih rukopisa manastira Hilandara. Paleografski album.* Belgrade.

Bulatova, R. V. (1975). *Staroserbskaja glagol'naja akcentuacija.* Moscow.

— (1977). 'K izučeniju sistem nadstročnyx znakov v drevneserbskix rukopisjax', in Dimitrije Bogdanović, Biljana Jovanović-Stipčević, and Đorđe Trifunović (eds.), *Zbornik Vladimira Mošina.* Belgrade, 113-16.

— (1981). 'Akcentuacija neproizvodnyx imen suščestvitel'nyx *a*-osnov v drevneserbskix rukopisjax XIV-XVI vv.', in *Slavjanskoe i balkanskoe jazyko-znanie. Problemy morfonologii.* Moscow, 60-89.

— (1989). 'Akcentuacija neproizvodnyx suščestvitel'nyx -*o*-osnov mužskogo roda v staroserbskom', in *Istoričeskaja akcentologija i sravnitel'no-istoričeskij metod.* Moscow, 204-24.

Daničić, Đura (1864). *Nikoljsko jevanđelje.* Belgrade.

Davies, Anna Morpurgo (1987), 'Folk-linguistics and the Greek word', in George Cardona and Norman H. Zide (eds.), *Festschrift for Henry Hoenigswald.* Tübingen, 263-80.

Đorđić, Petar (1971). *Istorija srpske ćirilice.* Belgrade.

Dybo, V. A. (1981). *Slavjanskaja akcentologija.* Moscow.

Džurova, Aksinija, Stančev, Krasimir, and Japundžić, Marko (1985). *Opis na slavjanskite răkopisi văv Vatikanskata biblioteka.* Sofia.

Grickat, Irina (1961-2). 'Divoševo jevanđelje'. *Južnoslovenski filolog*, 25: 227-91.

Grković-Mejdžor, Jasmina (1993). *Jezik 'Psaltira' iz štamparije Crnojevića*. Podgorica.

Hinrichs, Jan Paul (1985). *Zum Akzent im Mittelbulgarischen*. Amsterdam.

Ivanova-Mavrodinova, Vera, and Džurova, Aksinija (eds.) (1981). *Asemanievoto evangelie. Faksimilno izdanie*. Sofia.

Ivić, Pavle, and Jerković, Vera (1981). *Pravopis srpskohrvatskih ćirilskih povelja i pisama XII i XIII veka*. Novi Sad.

—, and Jerković, Vera (1982). *Paleografski opis i pravopis Dečanskih hrisovulja*. Novi Sad.

Jagić, Vatroslav (ed.) (1890). *Svetostefanski hrisovulj kralja Stefana Urosa II Milutina*. Vienna.

Jerković, Vera (1975). *Paleografska i jezička ispitivanja o Čajničkom jevanđelju*. Novi Sad.

Karskij, E. F. (1928). *Slavjanskaja kirillovskaja paleografija*. Leningrad.

Kos, Janko, *et al.* (ed.) (1993). *Brižinski spomeniki. Znanstvenokritična izdaja*. Ljubljana.

Kuna, Herta (1977). 'Neke grafijske osobine bosanskih srednjovjekovnih kodeksa u odnosu prema staroslavenskoj glagoljskoj grafijskoj tradiciji', in *Slovansko jezikoslovje. Nahtigalov zbornik*. Ljubljana, 153-66.

— (ed.) (1986). *Zbornik Hvala krstjanina; transkripcija i komentar*. Sarajevo.

Lavrov, P. A. (1914). *Paleografičeskoe obozrenie kirillovskogo pis'ma* (Enciklopedija slavjanskoj filologii, 4.1). St Petersburg.

MacRobert, C. M. (1989). 'Two for the price of one: the Psalter MS Peć 68'. *Oxford Slavonic Papers* NS, 22: 1-33.

Martinović, Dušan (ed.) (1986). *Psaltir s posljedovanjem Đurđa Crnojevića 1494*. Cetinje.

Medaković, Dejan (1958). *Grafika srpskih štampanih knjiga XV-XVII veka*. Belgrade.

Mošin, Vladimir (1966). *Paleografski albom na južnoslovenskoto kirilsko pismo*. Skopje.

Nedeljković, Olga (1967). 'Znaki udarenij v serbskix doresavskix rukopisjax'. *Slavia*, 36: 24-32.

Novaković, Stefan (1877). 'Akcenti štampanih srpsko-slovenskih knjiga crno-gorskih i mletačkih'. *Glasnik srpskog učenog društva*, 44.

Obnorskij, S. P., and Barxudarov, S. G. (1952). *Xristomatija po istorii russkogo jazyka* (2nd edn.). Moscow.

Osipov, B. I. (1992). *Istorija russkoj orfografii i punktuacii*. Novosibirsk.

Parkes, M. B. (1992). *Pause and Effect: An Introduction to the History of Punctuation in the West*. Aldershot.

Pavlović, M. (1958). 'Sistema udarenija Radoja D'jaka, pisca gramoty Bana Kulina', in *IV Meždunarodnyj s"ezd slavistov. Slavjanskaja filologija, II*. Moscow, 75-95.

Rodić, Nikola, and Jovanović, Gordana (1986). *Miroslavljevo Jevanđelje. Kritično izdanje*. Belgrade.

Saenger, Paul (1997). *Space between Words: the Origins of Silent Reading*. Stanford.

Schaeken, Joos (1987). *Die Kiever Blätter*. Amsterdam.

Šidak, Jaroslav (1955). 'Kopitarovo bosansko evanđelje u sklopu pitanja "crkve bosanske"'. *Slovo*, 4-5: 47-63.

Speranski, Mixail (1906). 'Mostarskoe (Manojlovo) evangelie'. *Russkij filologičeskij vestnik*, 54: 1-95.

Stang, Christian (1978). 'Zur Frage der Betonung univerbierter Bildungen im Russischen', in Henrik Birnbaum *et al.* (ed.), *Studia Linguistica Alexandro Vasilii filio Issatschenko a collegis amicisque oblata*. Lisse, 397-403.

Štavljanin-Đorđević, Ljubica, Grozdanović-Paić, Miroslava, and Cernić, Lucija (1991). *Opis ćirilskih rukopisa Narodne Biblioteke Srbije*, *1. Paleografski album*. Belgrade.

Störmer, Olaf (1987). *Die altrussischen Handschriften liturgischer Gesänge in sematischer Notation als Hilfsmittel der slavischen Akzentologie* (Slavistische Beiträge, 20). Munich.

Tachiaos, A.-E. N. (1981). *The Slavonic Manuscripts of Saint Panteleimon Monastery (Rossikon) on Mount Athos*. Thessaloniki/Los Angeles.

Trifunović, Đorđe (2001). *Ka počecima srpske pismenosti*. Belgrade.

Trubetzkoy, N. S. (1954/1968). *Altkirchenslavische Grammatik. Schrift-, Laut- und Formensystem*. Vienna/Graz.

Vajs, Josef (1932). *Rukovět' hlaholské paleografie*. Prague.

Večerka, Radoslav (1989). *Altkirchenslavische (altbulgarische) Syntax. I. Die lineare Satzorganisation*. Freiburg.

Vrana, Josip (1967). *Vukanovo Evanđelje*. Belgrade.

Zaliznjak, A. A. (1985). *Ot praslavjanskoj akcentuacii k russkoj*. Moscow.

— (1995). *Drevnenovgorodskij dialekt*. Moscow.

From the etymology workshop

Work in progress on the Oxford English Dictionary

*Philip Durkin, Jane McCauley, Tania Styles, Samantha Schad,
Neil Fulton, Anthony Esposito, Joanna Tulloch*

1. Introduction: the third edition of the OED

The *Oxford English Dictionary* (*OED*) is presently undergoing its first ever full, top-to-bottom, revision. The full text of the Dictionary, first published between 1884 and 1928, and supplemented in 1933 and again between 1972 and 1986, was brought together in a single sequence in the second edition of 1989 (henceforth *OED2*). This (together with material from the *OED Additions Series* volumes) has now been published online at <http://www.oed.com> where the existing text is now being supplemented with quarterly updates of both revised and newly added material in alphabetical sequence (so far approximately 8000 entries between M and MESYLATION), as well as with entries for newly added items across the whole alphabet, which together constitute the first instalments of the third edition of the Dictionary (henceforth *OED3*).

No dictionary, however extensive, ever allows its editors enough space to justify their decisions in full, nor can it account for all its sources, or properly distinguish between what is new and what is derivative. The reader must take a number of conclusions on trust, and often misses some of the more general developments. The pieces below are intended to illustrate something of the range and scope of work presently being done on etymologies for the third edition of the *OED*.[1] The examples are arranged thematically, and have been chosen by members of the *OED*'s etymology team with the following aims in mind. Firstly, they offer concrete illustrations of different aspects of the etymological work being done for *OED3*. Secondly, and equally importantly, they have in many cases also been selected for more personal reasons, as being dictionary entries that it has been particularly satisfying to revise. It is hoped that the reader may come to share in some of the sheer satisfaction of working through the *OED*'s etymologies, while at the same time seeing some of the recurrent features which run through the revised text.   *Philip Durkin*

---

[1] They are not intended to give a full overview of the etymological policy and procedures of *OED3*, for a fuller account of which see Durkin (1999).

2. *Mixed transmission*

Our ability to present an accurate account of how words enter the English language relies to a great extent on the quantity and quality of source material available for each donor language. While other sections of this paper will show the research which has been carried out for *OED3* on non-dictionary sources for foreign languages (such as seeking out the originals of translations quoted within the Dictionary, or the use of electronic corpora, or work-titles in library catalogues, etc., as sources), the first recourse for researching a loan into English is usually a foreign-language dictionary.[2]

One of the areas where it is possible to observe the effect of the growth in foreign-language dictionary materials available to *OED3* editors is the increased number of entries which present a mixed or multiple etymology, or acknowledge a previously unremarked element of uncertainty in accounting for a word's transmission into English. A great many of these entries involve the acknowledgement of a potential Romance transmission route in an etymology which was previously considered to have been entirely from Latin, or vice versa.

Such entries tend to fall into the following categories:

(1) Entries previously etymologized as reflecting one direct route into the language, where new information has now shown that a mixed origin is more likely, e.g. MANILLE *n*. (the second highest scoring trump or honour in quadrille, ombre, etc., and also the name of a cards game):

> **manille** *n*., *OED2*, forms list and etymology:
>> Forms: 7 **mal(l)illio**, 9 **malilla**, 8 **manil(l)io**, **manill**, 8- **manille**.
>> [Corruptly a. Sp. *malilla*, dim. of *mala* used in the same sense (prob. fem. of *malo* bad).]
>
> **manille** *n*., *OED3*, forms list and first part of etymology:
>> Forms: α. 16 **malillio**, **mallillio**, 18 **malilla**. β. 17 **manilio**, **manill**, **manillio**, 17-18 **menille**, 17- **manille**, 18 **manilla**, 19- **menel**.
>> [Partly < French *manille* (1696 in sense 1; only attested from 1883 in sense 2, and from 1893 denoting the ten of each suit; 1660 in a Spanish-French dictionary in forms *malille* and *menille* glossed as nine of diamonds), and partly < its etymon Spanish *malilla* the second most valuable card in certain games (1604) < *mala*, spec. use in same sense of the feminine of *malo* bad (< classical Latin *malus*: see MAL-) + -*illa* -ILLA. The French form *manille* (and prob. hence the English β forms) prob. arose by dissimilation, perh. influenced also by *main* hand; cf. also Catalan *manille* < Spanish…]

---

[2] Space does not permit here to give even an overview of the foreign-language dictionary resources routinely drawn upon for *OED3*. For a preliminary selective listing see Durkin (1999).

In *OED2* the current English form in *-n-* was explained only as a corruption of the Spanish form. In *OED3*, as a result of French evidence provided by the *Trésor de la Langue Française* it is now possible to see that these *-n-* forms are likely to have resulted from French influence.

(2) Entries previously etymologized as Romance loans which now acknowledge a Latin element:

mantlet *n.*, *OED3*, first part of etymology:

[Partly < post-classical Latin *mantellettum*, *mantilettum* short mantle (from 1316 in British sources), caparison (1316 in a British source), movable shelter (from 1374 in British sources), mantelletta (as *mantelletum*, in an undated source cited in Du Cange), and partly from its probable etymon Anglo-Norman and Old French, Middle French *mantelet* short mantle (*c*1140; 1553 in Middle French in sense 2a, although 14th cent. in Old Occitan; 1679 in French in sense 1c) < *mantel* MANTLE *n.* + *-et* -ET[1]. Cf. Italian *mantelletto* (14th cent.).]

(3) Entries previously etymologized as Latin loans which now acknowledge a Romance element:

macaronic *a.*, *OED3*, first part of etymology:

[< Middle French, French *macaronique* (1552 *vers macaronicques* Rabelais) or its etymon post-classical Latin *macaronicus* (see below) < Italian †*macaroni* MACARONI *n.* + *-icus* -IC. Cf. Italian †*macaronico*, *maccheronico* (1634).]

The increased attention paid to senses and dates in these foreign language dictionaries has also, on occasion, provided us with the information necessary to re-evaluate a word's loan status. In some instances (e.g. MALEFICIOUS *a.*, MATRICULARY *n.*) a word previously believed to be a loan due to its formal equivalence to an item in a Romance or classical language has now been re-assessed as more likely to be an English formation.

The inclusion of these mixed or multiple etymologies is by no means an *OED3* innovation, but access to more detailed foreign-language dictionaries often means that *OED3* etymologies now have the necessary evidence to support what could previously only be given as a suggestion. One small example of this can be seen at MECHANIC *a.* and *n.* While *OED2* cited this word as a Latin loan, it also acknowledged that there was probably some Romance influence in its early transmission:

> **mechanic** *a*. and *n*., *OED2*, extract from etymology:
>
> [ad L. *mēchanic-us*... Cf. F. *mécanique* (from 14th c.: perh. the source in early instances)...]

With the more detailed sense and date information provided by *TLF*, *OED3* is now able to include this as an item with a definite mixed origin:

> **mechanic** *a*. and *n*., *OED3*, first part of etymology:
>
> [Partly < Middle French *mécanique* (adjective) characterized by use of tools and the hands (*c*1265 in Old French), (noun) manual worker, artisan (13th cent. in Old French), and partly < its etymon classical Latin *mēchanicus* (adjective) mechanical, concerning machines, (noun) engineer < ancient Greek μηχανικός resourceful, relating to machines (also as noun in sense 'engineer' in Hellenistic Greek) < μηχανή MACHINE *n*. + -ικός -IC.

With more detailed foreign-language dictionary sources to hand, it is thus possible to evaluate more clearly the probable route by which a word has entered the English language. In this way, a word's transmission can be more accurately reflected in these (relatively) more complex etymologies.

<div align="right">

Jane McCauley

</div>

3. Tracking down foreign-language documentation: wine names

Vital to the task of composing and revising the etymologies of English words in *OED3* are the historical dictionaries of other languages, which are routinely used to establish forms, senses, and dates of first record for foreign-language words. Advances in the historical lexicography of all the European languages have been considerable since the first edition of the Dictionary, and as has been shown above, the improved documentation now available in these works has enabled us to give a more accurate, full, and sophisticated account of the transmission into and the development within English of a considerable number of loanwords. Conversely, research into the history of English loanwords sometimes yields information about the donor language which is not available in the relevant foreign-language dictionaries. This is demonstrated well by a particularly fruitful source for evidence of linguistic and cultural contact: the vocabulary of wine and winemaking. The examples discussed below all belong to this group.

In the majority of cases, the data available in foreign-language dictionaries confirm that a loanword has entered English by the expected route. Consider, for example, the entry for MÉDOC *n*.:

médoc *n.*, *OED3*, etymology:

> [< French *médoc* (1789) < *Médoc*, the name of a wine region in south-western France, on the left bank of the Gironde River estuary, north-west of Bordeaux.]

Here, the *Trésor de la Langue Française* records French *médoc* as the name of a wine in 1789, nearly half a century earlier than the first use of the word in English in this sense (1824 in the compound *médoc wine*, 1833 as simplex), providing a chronology which supports *OED2*'s assertion that the English name for the wine is borrowed direct from the French. It goes on to confirm the etymology for the French word from the name of the wine-producing region. By contrast *merlot*, the name of another French wine (and the grape from which it is made), is attested in English from 1825, but the earliest attestation given by French dictionaries is 1861 (in *TLF*), which would appear to turn the intuitive course of transmission on its head. However, the French text which is translated in the 1985 quotation discusses a French document of 1784 which mentions the 'Merlau' grape. The word was therefore in use in French almost a century before *TLF*'s first attestation, and available to be borrowed into English by the time of OED3's first quotation:

> **merlot** *n.*, *OED3*, forms, etymology, definition, and quotations:
>
> Forms: 18 **Murleau**, 18- **Merlot**, (19- *hist.* **Merlau**). Also with lower-case initial.
>
> [< French *merlot* (app. late 18th cent. as *merlau*) < *merle* blackbird (see MERLE *n.*[1]) + *-ot* -OT[1], with allusion to the colour of the grapes.]
>
> A variety of vine yielding black grapes used in winemaking, widely grown throughout France, (esp. around Bordeaux), in California, and in other temperate regions; the grape of this variety; the soft-textured, distinctively fruit-flavoured red wine made from this grape. Also ***Merlot noir***.
>
>   **1825** J. BUSBY *Treat. Culture Vine* ii. 71 The *murleau*. This variety announces much vigour, by the strength of its wood. **1833** C. REDDING *Hist. Mod. Wines* v. 141 The vine plants most cultivated in the canton of Bourg are the *merlot*, the *carminet*, the *mancin*, the *teinturier*, the *petit chalosse noire*, and..the *prolongeau*. **1888** *Encycl. Brit.* XXIV. 604/2 The vines of the Cabernet species..are not so greatly used as the Merlot, which is very productive, and not so liable to attacks from *Oidium*. **1926** P. MORTON SHAND *Bk. Wine* v. 88 The grapes are cabernet, merlot, malbec; and one vineyard has a chateau built by Richelieu. **1978** *Amer. Poetry Rev.* Nov.-Dec. 22/3 Two or three empty bottles of Merlot, Avant-garde of the gallons that are to come. **1985** H. COLEMAN & R. MAXWELL tr. H. Enjalbert *Great Bordeaux Wines* II. 223 Only two, including the 'Merlau' are excellent. **2000** *Wine* May 17/3 We believe that La Capitana offers the perfect soil..to produce both the Merlot and Sangiovese grapes that comprise the Lucente blend.

Not infrequently we find that documentation is more plentiful for English than for the likely donor language with more significant results. A case in point is MARGAUX *n.*:

> **Margaux** *n.*, *OED3*, etymology, definition, and earliest quotations:
>
> [< French *Margaux*, the name of a commune in the department of Gironde, France. French *Margaux* may refer to the wine of Château Margaux, or to wine of the commune of Margaux, in which a number of wine-producing propriétaires are situated, or to wines which are entitled to bear the name Margaux under the legislation regulating the Appellations d'Origine Contrôlées (including many but not all of the red wines from the commune of Margaux, and from the neighbouring communes of Cantenac, Soussans, Arsac, and Labarde).]
>
> A red Bordeaux wine of the Haut-Médoc produced in Margaux, *esp.* that produced at Château Margaux (more fully ***Château Margaux***).
>
> **1705** *London Gaz.* No. 4128/4, 230 Hogsheads of new Pontac and Margose Wine. **1725** L. WELSTED *Oikographia* 12 Nor Margou, stor'd in Priestly Cells, That on the Palate grateful dwells. **1734** T. SHERIDAN *Let. to Swift* 16 Aug., in *Swift's Wks.* (1841) II. 724, I drink right French margose. **1793** T. JEFFERSON *Memorandum* 20 Feb. in *Papers* (1997) XXVII. 843 Bordeaux red wines. There are four crops of them more famous than all the rest. These are Chateau-Margau, [etc.]. *a***1845** R. H. BARHAM *Ingoldsby Legends* (1847) 3rd Ser. 186 Chambertin, Chateau Margaux, La Rose, and Lafitte.

No corresponding noun is recorded in any of the major French dictionaries, in any sense; corpora such as the *Frantext* corpus[3] give us *Château-Margaux* denoting the wine, but only from the mid nineteenth century. In the absence of evidence for a corresponding French word as a proximate source, the etymology of the English name for the wine must be given directly from the French place name. It is difficult to believe that this English name for a wine produced in the heart of the Médoc region of France and referred to by Sheridan in 1734 as 'right French margose' does not reflect similar use in French, but the available data does not enable us to make such an assertion.   *Tania Styles*

4. Etymology and translated sources

A number of lemmata in the *OED* are first attested in texts which are translations. It is often fruitful from an etymological point of view to investigate

---

[3] Frantext, prepared by the Institut National de la Langue Française, <http://zeus.inalf.cnrs.fr/frantext.htm>.

the sources from which these translations were made. A word first recorded in a translation from, say, Spanish or Italian, is quite likely to have been borrowed from that language or at least modelled on the corresponding word in the foreign-language source. We have done a great deal of work in this area, and built up a large in-house working database (compiled from such sources as editions, library catalogues, the *Dictionary of National Biography*, and secondary literature) of bibliographical information on texts which are translations, together with the translated sources. Ideally, it is desirable to consult the texts side by side, the original with the translation, but it is not always possible to locate both in the same library. Through our research we have been able in many instances to antedate foreign-language forms recorded in dictionaries, or to provide evidence for words omitted from dictionaries, for example as being too rare or transient, or otherwise failing the inclusion criteria of a particular lexicographical tradition (as for example much scientific vocabulary, or formations from proper nouns).

Translators work with widely differing degrees of accuracy and fidelity to their original source. Obviously, it depends to some extent whether the work concerned is a technical or a literary one, and whether the translator is aiming to render the substance, spirit, or actual words of the base text. Biblical translations, for example, tend to be strictly verbatim, whereas translations of literary works vary from the closely literal to versions so loose they may barely be called translations at all. Moreover, translations are not always made from the ultimate foreign-language source, but may be made from an intermediary translation. This is often the case with languages such as ancient Greek and Arabic, less widely known. For example, MANTINEAN *n.* and *a.* is first recorded in a translation of Thucydides, but the translator, who knew no Greek, was dependent on a French version:[4]

> **Mantinean** *n.* and *a.*, *OED3*, etymology and first part of first sense:
> [< classical Latin *Mantinēa* (transcription of ancient Greek Μαντίνεια) Mantinea, the name of a city in Arcadia, ancient Greece + -AN after Middle French *Mantinien*, noun (1545 in the passage translated in the quot. 1550). Cf. post-classical Latin *Mantineus*, noun (*a*1457 in de Valla's translation of Thucydides), ancient Greek Μαντινεύς, noun, Μαντινικός, adjective. The form in *-eian* is after the Greek form.]
> **A.** *n.* A native or inhabitant of the ancient city of Mantinea.

---

[4] For full bibliographical details of translations and their immediate models see the main bibliography below.

> **1550** T. NICOLLS tr. Thucydides *Peloponnesian War* V. fol. 145 That the Argiues shulde make allyance wyth the Lacedemonyans, forsakinge the same, whyche they had concluded wyth the Athenyans, the Mantynyans and the Elyans. …

Then there are some red herrings. The first quotation in *OED2* for MEDITERRANEANIZE *v*. 'to make Mediterranean in character or attributes' is:

> 1947 N. CARDUS *Autobiogr.* III. 244 He mediterraneanises Wagner, to use Nietzsche's term.

This might suggest that the term is taken from German, and in fact *OED2* documents an 1896 translation of Nietzsche incorporating the French phrase 'Il faut méditerraniser la musique'. Investigation reveals that Nietzsche used the phrase in French in the original German work, and gave no German equivalent:

> **Mediterraneanize** *v*., *OED3*, etymology:
> < MEDITERRANEAN *a*. + -IZE. Cf. earlier MEDITERRANEANIZED *a*. and MEDITERRANEAN-IZATION *n*., and also French *mediterraniser*, attested in an English context (and in a German context in the passage translated) in:
> **1896** T. COMMON tr. F. Nietzsche *Case of Wagner* iii. 9 *Il faut méditerraniser la musique*: I have no reasons for using this formula. The return to nature, to health, to gaiety, to youth, and to *virtue*!

One area in which research of this nature has proved particularly fruitful is with regard to scientific and technical vocabulary. The name of the chemical substance *melampyrite* is first recorded in a text-book translated from German:

> **melampyrite** *n*., *OED3*, etymology, definition, and quotations:
> [< German *Melampyrit* (1861 in the passage translated in quot. 1865) < scientific Latin *Melampyrum* MELAMPYRE *n*. + German *-it* -ITE[1]. Cf. French *melampyrite* (1873).]
> = DULCITOL *n*.
> **1865** H. WATTS tr. L. Gmelin *Handbk. Chem.* XV. 389 Melampyrite. **1882** H. WATTS *Dict. Chem.* (new ed.) II. 348 Quite recently Gilmer has shown that melampyrite..is really identical in composition and properties with dulcite.

Translations may also offer evidence of use for the original language. German dictionaries, for instance, tend to provide etymological information for only the commonest scientific and technical words. Consequently the investigation of translations and their sources can provide information as valuable for German as for English.   *Samantha Schad*

5. Pronunciation history, spellings, and slang etymology

One field in which the quality of available documentation has improved beyond compare since the production of the first edition of the Dictionary, allowing the current editors to say much more than their predecessors were able to, is the study of pronunciation and its changes through history. One very obvious example of this can be found in the entry for Maya *n.*[2] and *a.* (the ancient American people):

> **Maya** *n.*[2] and *a.*, *OED3*, pronunciation and note from etymology section:
> *Brit.* /ˈmʌɪə/, *U.S.* /ˈmaɪə/.
> The pronunciation of this word has been subject to considerable variation. *Cent. Dict.* (1890) gives the pronunciation as /ˈmɑːjɑː/. The word is first found in Webster in 1911 with the pronunciation /ˈmɑjɑ/; the pronunciation /ˈmaɪ(j)ə/ appears in the 1954 ed., and becomes the primary pronunciation in the 1961 ed., when a third variant /ˈmeɪ(j)ə/ is also recorded. *O.E.D. Suppl.* (1976) gives the pronunciation as (māˑ·yă, məi·(y)ă, mēˑ·ă) /ˈmɑːjə/ /ˈmaɪ(j)ə/ /ˈmeɪə/. Dicts. of the close of the 20th cent. generally give the pronunciation as *Brit.* /ˈmʌɪə/, *U.S.* /ˈmaɪə/.

Here we have a very clear illustration of the variability in pronunciation typically shown by words of foreign origin when they first enter the language, and the process by which a consensus slowly emerges as to which of the possible pronunciations will become dominant. Because all of this process has taken place within the comparatively recent past, the account can be backed up with considerable detail.

There are other areas too in which the increased quality and accessibility of documented sources allows us to answer questions which might previously have remained mysterious. A personal favourite, illustrating the progress which has been made in just the last few years, is our discovery of how the *Mary Celeste* came to change its spelling. In 1997, *OED Additions Series* vol. 3 noted that the name of the 'Mary Celeste', an abandoned American cargo ship, had come to be used allusively to refer to any inexplicably deserted place, but no explanation could be offered of why the word is now almost always spelled *Marie* (rather than *Mary*) *Celeste*. *OED3* is able to complete the account:

> **Marie Celeste** *n.*, *OED3*, forms list and etymology:
> Forms: 19- **Marie Celeste**, **Mary Celeste**.
> [< the name of the *Mary Celeste*, an American cargo ship which in December 1872 was found mysteriously abandoned in the North Atlantic with sails set. In form *Marie Celeste* reflecting the spelling used in the highly fictionalized account by Arthur Conan

> Doyle, *J. Habakuk Jephson's Statement*, published pseudonymously in the *Cornhill Mag.* of Jan. 1884; before this the mystery had received little public attention.]

In many ways this is typical of the process of dictionary making. After all, it was obvious enough that *some* person, at *some* time, must have made a misspelling, which later became widespread: the lexicographer's satisfaction lies in establishing that it was *this* person, at *this* time.

But it would be unfair to suggest that a diligent search of the relevant records will inevitably allow the lexicographer to come up with the right answer. With some questions, the further one delves into the archives, the more obvious it becomes that no satisfactory answer is ever likely to emerge. The most spectacular example of this phenomenon which we have recently encountered is that of the phrase *the full monty*:

> **full monty** *n*. (and *a*.), *OED3*, etymology:
> [< FULL *a.* + *monty* (origin unknown).
> Many theories are proposed as to the origin of this phrase, but none of them is supported by reliable historical evidence. Perh. the most plausible is that it is from a colloquial shortening of the name of *Montague* Maurice Burton (1885-1952), men's tailor, and referred originally to the purchase of a complete three-piece suit. Also popular but unsubstantiated is the belief that the phrase is somehow derived from *Monty*, the nickname of Field Marshal Bernard Law Montgomery (1887-1976). However, the sheer variety of often vague, purely anecdotal, and mutually contradictory explanations for the connection — ranging from his wartime briefing style to his breakfasting habits — renders this less credible. Other suggestions, including references to MONTY *n*. and MONTE *n*.[1], are still more speculative.]

The search for the origin of this phrase seems likely to remain intractable. On most occasions when this is the case, it is because there is simply no clear idea of where a word or phrase might come from. This case is unusual because there are so many different ideas. Over the course of two or three years' research, including appeals to the public for information, which met with a typically generous response, we were able to collect a total of seventeen independent theories of the origin of the phrase, many of them with a number of variations. Unfortunately, the one thing which all of these theories had in common was that there was no clear evidence to support them. Indeed, the most frustrating aspect of work on this phrase has been the contrast between the very common impression among our correspondents that they recall its use many years ago, and our complete inability to trace an example of it in print from any earlier than 1985. In these circumstances, the lexicographer's primary responsibility is not to

misrepresent the facts, and we make it plain from the start that we make no claim to knowing the origin of the phrase. Only then do we allow ourselves the luxury of discussing one or two of those possible explanations which we think might be worth pursuing further. *Neil Fulton*

6. Taking a new look at Old English

The preparation of *OED3* affords a welcome opportunity for the complete overhaul of Old English material in *OED*. Although the date 1150, chosen by Murray, remains the *terminus a quo* for the *OED*'s documentation of the history of the English Language, in practice this absolves *OED* only from the need to record words which did not survive in English later than this date; all words which do continue in use after 1150 are given full documention back to earliest times. The fruits of almost 100 years of research in Old English can now be brought to bear on the revision of *OED*. The text of all Old English quotations is being verified (a surprising number of inaccuracies has been discovered in the process); bibliographical details are also being verified, and wherever possible, conversion is being made to more recent editions; lists of Old English variant forms are being made more comprehensive. Reference to Old English within etymologies is also being greatly extended. This may be done on the basis of a wide variety of criteria, including the following:

- *Phonological*, e.g. at MANSLAUGHT *n.*, there is a long discussion about the importance of the presence or absence of *i*-mutation in Old English to explain the divergence in the form history; at MARE *n.*[1], the merging of two distinct OE words (*mearh* and *mīre*) is explained; at MAY *v.*[1], the complicated Old English phonology is treated in much greater detail than before;
- *Syntactical*, e.g. at MAN *pron.*, the precise grammatical usage of the word is explained; at MERRY *a.*, the development of the modern use of the word with a personal referent is traced back to certain Old English constructions;
- *Orthographic*, e.g. at MAN *n.*[1], the relative frequency of *mann* and *monn* spellings is discussed;
- *Lexical*, e.g. at MAKE *v.*[1], there is a note about the precise usage of OE *macian*, and its relative infrequency compared to *dōn* and other verbs; at MARTER *n.*[1], information is given about the Old English word (*mearþ*) that was superseded by the Old French borrowing *marter*, and its survival in the compound FOUMART *n.* and as an element in place names;
- *Encyclopaedic*, e.g. at MACCABEE *n.*, it is noted that the name in its post-classical Latin form *Machabeus* occurs in Old English biblical translations; at MACEDONIAN *n.*[1] and *a.*[1] and MAURITANIAN *a.* and *n.*, the occurrence of the

place names *Macedonia* and *Mauritania* in English from Old English onwards is noted.

I will conclude with one or two examples which illustrate what is arguably one of the most gratifying outcomes for the historical lexicographer: the discovery that a word or sense is older than was thought. Work on Old English material for *OED3* has already turned up a fine crop of antedatings, sometimes of senses of words (e.g. MEAL *n.*[1], sense 2a, MEAT *n.*, sense 3), sometimes of words themselves. Such antedatings often necessitate quite radical changes to the etymology of the words in question, e.g. MERRY *adv.* is now a separate headword (where before it was subsumed in the entry for MERRY *a.*), with the etymology containing a discussion of mutated and unmutated adjective and adverb pairs in Old English:

> **merry** *adv.*, *OED3*, etymology:
>
> [Partly < an adverbial derivative of the Germanic base of MERRY *a.*, and partly (in later use) directly < MERRY *a.*
>
> The Old English form *murge* has an unmutated root vowel (in contrast with the mutated vowel of the adjective *myrge*). This pattern is repeated in several Old English adverb/adjective pairs (e.g. *ange* anxiously and *enge* anxious, *clāne* cleanly and *clǣne* clean, *swōte* SOOT *adv.* and *swēte* SWEET *a.*), and prob. results from the fact that all of these words were orig. *u*-stems, which in the adjectival forms have gone over to the *ja*-stem declension; analogical forms (cf. *myrge*, adverb) are common.]

The earliest attestation of the word is pushed back by over 150 years from OED2's date of *c.* 1220:

> **merry** *adv.*, *OED3*, earliest examples:
>
> **OE** *Homily* (Hatton 113) in A. S. Napier *Wulfstan* (1883) 152 Hu myrge he sang mid þam munecum symle. **lOE** ÆLFRED tr. Augustine *Soliloquies* (Vitell.) Pref. 47 And fegerne tun timbrian,..and þær murge and softe mid mæge on-eardian. *c1175 Cnut's Song* (Trin. Cambr.) 1 in E. O. Blake *Liber Eliensis* 153 Merie sungen ∂e muneches binnen Ely ∂a Cnut ching reu ∂er by.

For the adjective *melch* ('mellow, soft, tender, etc.', now antedated from the Middle English period to Old English), a much stronger case is made for ultimate derivation from the Indo-European base meaning 'honey', and the probable influence of the Old English word on other Germanic languages is noted:

melsh, melch *a*., *OED2*, etymology:

[Perh. repr. OE. *męlsc, mylsc*, **mielisc* mellow (in *melsc æppla*, mellow apples), ?cogn. w. Goth. (*ga*)*malwjan* to crush. The OE. word seems to have been confused with *milisc* honeyed, cogn. w. Goth. *miliþ* honey. Cf. MULSH.]

melch *a*., *OED3*, etymology:

[Origin uncertain: prob. ult. < the Indo-European base of MELL *n.*[2] (cf. the first element of MILDEW *n.*) + -ISH[1], a derivation which accords well with the sense of the word in Old English, but in later use app. influenced by and partly remodelled after an adjective ult. < the Indo-European base of MEAL *n.*[1] and represented by MULCH *a*.

Cf. Old High German *milsca* (in glosses) mixed drink of honey and wine, late Old Icelandic *milska* mixed drink of mead and ale, Norwegian (Nynorsk) *mylske* kind of sweet cheese, mixture of cheese and meal, Old Swedish *mylska*, *mölska* strong and sweet ale (Swedish regional *mölska*), Danish regional *melske* mixed drink of mead and ale, all perh. ult. borrowings from Old English; cf. also Old English *milscian* to sweeten a drink with honey (also twice glossing post-classical Latin *mitescere* to become soft and sweet), Old High German *milsken* (in glosses) to sweeten wine, esp. with honey, late Old Icelandic *milska* to mix a drink.

In Old English the compound *sūrmilisc* sourish-sweet, is also attested, cf.:

eOE *Bald's Leechbk.* II. i. 176 Forþon þæm mannum deah þæt him mon on fruman þa mettas gife þe celunge & strangunge mægen hæbben swa swa beoþ æppla nales to swete ealles ac surmelsce.

In forms in -*i*- in sense 3 perh. remodelled after MILCE *n.*: see *milce-hearted* adj. s.v.]

Full documentation is given of the Old English sense 'sweet, honeyed', which is only cursorily mentioned in *OED2*:

melch *a*., *OED3*, definition and earliest examples:

†**1.** Sweet; (of a drink) honeyed, mulled; (of a tree) producing sweet fruit. *Obs.*

eOE *Épinal Gloss.* 34 *Melarium*, milsc apuldr. **eOE** *Bald's Leechbk.* I. xlii. 108 Wyrc him þonne stilne drenc of ompran on wine & on wætre & on þam baðe gehwilce morgene drince mylsce drincan.

To give a final example, at MERE *a.*[2] ('pure, etc.'; now also antedated from the Middle English period to Old English), much fuller evidence is given (in the quotation paragraphs) for the word's borrowing (from post-classical Latin) during the Old English period (contrast this with the rather tentative introduction into the *OED2* etymology of only a fraction of the Old English evidence). The *OED3* etymology traces in detail two distinct layers of borrowing: an earlier borrowing into Old English from post-classical Latin, reinforced by a later borrowing into Middle English from Anglo-Norman. In addition, much

important semantic information appears in the etymology for the first time, citing post-classical Latin and Old French models for specific English collocations:

> **mere** *a.*[2], *OED3*, extracts from etymology:
>> [Prob. partly (esp. in early use) < a post-classical Latin form (with characteristic vulgar Latin lengthening of vowels in open syllables) of classical Latin *merus* undiluted, unmixed, pure < the same Indo-European base as MERE *v.*[1], and partly (in Middle English) a reborrowing of its reflex Anglo-Norman *mer*, *meer*, *mier*, Middle French *mer* (*c*1100 in Old French as *mier*)… In classical Latin senses corresponding to 1a, 1b, and 4 are already found, as is the sense 'nothing more than, simple' (cf. sense 5). In post-classical Latin chiefly designating unmixed wine or unalloyed metals or coinage, and in various collocations... In Old French and Anglo-Norman found chiefly in sense 'pure' as applied to metals, esp. gold…]

Anthony Esposito

7. 'Exotic' loans and their etymologies

In addition to etymologies for words of Germanic, Romance, Italic, and Celtic origin and those formed on Classical models, *OED3* presents much new information about words from more exotic languages. For the purpose of this section of the paper, an 'exotic loan' is a lexical item ultimately borrowed from any language outside the groups listed above. So far approximately 400 such words and phrases in *OED2* have been revised and published in *OED3*; the published material also includes more than 200 new headwords of exotic origin. Between two rounds of editing in-house, all of this material has been submitted for comment to specialist language consultants, some of whom have carried out new research on the *OED*'s behalf. For the new edition it has been possible to provide both fuller and more up-to-date etymologies for many of the words included in *OED2*, and an equivalent etymological treatment of the new material.

The word *matchcoat* (a kind of coat formerly worn by North American Indians) illustrates both the updating of an existing etymology in the light of recent scholarship and the provision of a completely new entry for its previously supposed etymon. The *OED2* etymology was based on a first quotation in the form *matchco* from 1642, followed by quotations predominantly in the form *Matchcoat* or *Match-coat*:

matchcoat *n.*, *OED2*, etymology:

[Orig. *matchco*, prob. an American Indian word: cf. Odjibwa *matchigode* 'petticoat, woman's dress' (Baraga); afterwards corrupted by popular etymology, as if f. MATCH *n.*[1] or *v.*[1] + COAT *n.*]

In the revised entry, two earlier quotations have been added, one (from a map of Virginia, dated 1612) showing the plural form *matchcores* and the other (1633) the form *match coate.* The new etymology reflects this fresh information:

matchcoat *n.*, *OED3*, etymology:

[< Virginia Algonquian *matchkore* deerskin robe < Proto-Algonquian **mat-* (empty root) + **-ixto*θ- robe, blanket. The oldest recorded form *matchcore* (see quot. *1612) reflects the etymon; the predominant later form shows assimilation of the final syllable to COAT *n.* by folk etymology. Cf. MATACHIA *n.*

Despite the similarity in form, the influence of the distinct word **MACHICOTE *n.*, or its Ojibwa etymon, is unlikely.]

The new entry MACHICOTE *n.* (a skirt worn, usually as an underskirt, by native American women), with quotations 1791–1941 illustrating a variety of spellings, presents the etymology of the Ojibwa word originally cited in the etymology of MATCHCOAT *n.*:

machicote *n.*, *OED3*, first part of etymology:

[< Ojibwa *mačiko:te:nʔ* skirt < Proto-Algonquian **mat-* (empty root) + **-iko:tay* skirt + a diminutive element. Cf. French *machicoté*, *machicôté* (*a*1761)…]

The word *mazurka* (a Polish dance in triple time) also illustrates the results of the revision process; the form of the Polish etymon has been corrected, raising the question of how the Polish masculine *mazurek* came to be borrowed as *mazurka*. A comparison of the English evidence with the appearance of the word in other western European languages reveals that it was probably transmitted into English via German or Russian:

mazurka *n.*, *OED2*, etymology:

[a. Polish *mazurka* woman of the Polish province Mazovia. In Fr. *masurka*, *mazurka*, *-ourka*, *-urke*, Ger. *Masurka*.]

mazurka *n.*, *OED3*, etymology:

[< Polish *mazurka*, accusative or genitive singular of *mazurek* folk dance from Mazovia (perh. extracted in an oblique case < a phrase such as *tańczyć mazurka* to dance the mazurka) < *mazur* mazurka, inhabitant of the Polish province Mazovia (Polish

Mazowsze: see MAZOVIAN *n.* and *a.*) or the adjacent Masuria (German *Masuren*, Polish *Mazury*: see *MASURIAN *a.*) + -*ek*, diminutive suffix. Cf. Russian *mazurka* (1795), French †*mazourka* (1828), *mazurka* (1831), †*mazourke* (1840), German *Masurka* (1795), Czech *mazurka*.

The borrowing (in all the main European languages) of the form in -*a* in preference to the nominative singular is paralleled by the borrowing of the synonymous Polish *mazur* into German as *Masura* (1752). The language by which the word was transmitted from Polish to western European languages was either Russian (in which the feminine diminutive suffix is traditionally used to name local dances) or perh. more likely German, in which there is abundant attestation of various names for the dance (*Masura*, *Masure*, *Massurisch*, etc.) from the mid 18th cent. onwards.]

In the examples given above, the starting point for the work was an existing etymology in *OED2*. However, many of the exotic loans in the Dictionary were not originally given etymologies. A distinction was made between assimilated and unassimilated loans, depending upon the degree of adaptation of the word to English morphology and phonology. For the unassimilated loans, and particularly for certain categories of word such as ethnonyms, exotic species, and culture-specific concepts, the etymological information given might be limited to an abbreviated form of the name of the source language, or even to the formula '[Native name.]'. For *OED3*, the distinction between assimilated and unassimilated loans has been dropped in favour of more detailed information on spellings, plural forms, and pronunciation in English. An etymology has been constructed for each of these words, aiming to show the etymon (including any forms in intermediary languages), the form of the word in the standard modern orthography of the source language (or, in the case of languages using non-roman scripts, a transliteration representing current philological best practice), a gloss (if not identical with the meaning of the word in English), and derivational morphology in the source language. Historical notes, especially to explain the history of spellings or pronunciation of the word in English or the fuller context of the word in the source language, together with notes on uncertain or disputed etymologies, are also given in certain cases.

The following examples illustrate greatly expanded etymologies compared with *OED2*, where MAST *n.* [6] (a kind of yogurt made in Iran and other parts of the Middle East) previously had the etymology '[Pers.]', MATABELE *n.* and *a.* (a member of the Ndebele people, and in other historical uses) had only '[Native name.]', and MEIJI *n.* and *a.* (the period of rule of the Japanese emperor Mutsuhito) had '[Jap., 'enlightened government'.]'. In the case of MEIJI *n.* and *a.*, an important etymological link has been made with the Chinese counterpart *Ming*, a more established and familiar borrowing in English:

mast *n.*⁶, *OED3*, etymology:

[< Persian *māst* yogurt, cognate with Khotanese *māsta* coagulated, Baluchi *mastaγ*, prob. < the Indo-European base of Sanskrit *mastu* sour cream, Armenian *macun* sour milk…]

Matabele *n.* and *a.* , *OED3*, etymology:

[< Sotho *Matabele* (singular *Letebele*) < *ma-*, plural prefix + *-tebele*, term applied by the Sotho and Tswana peoples to invading Zulu regiments, either < *teba* to sink down (with reference to the Zulu battle tactic of crouching behind their large shields), or < *thebe* shield. Cf. NDEBELE *n.*]

Meiji *n.* and *a.*, *OED3*, etymology:

[< Japanese *Meiji*, lit. 'enlightened government' < *mei-* shining light (< the Middle Chinese base of Chinese *míng*: see MING *n.*²) + *ji* peace, rule (< Middle Chinese).

The emperor Mutsuhito (posthumously, emperor Meiji) selected the name Meiji for his reign on acceding to the throne at the age of 16. The element *Mei-* (prob. under the influence of Chinese *míng*) had formed part of five earlier era names, and *ji* had formed part of 20. The element *mei-* also occurs in *Kōmei*, the name of Mutsuhito's father.]

Some words whose etymologies were unknown to *OED2* have also been provided with explanations, for example **mellah** (the Jewish quarter in certain North African towns):

mellah *n.*, *OED3*, etymology:

[< Moroccan Arabic *mallāḥ* Jewish quarter, generic use of *al-Mallāḥ*, lit. 'the saline area' (< *malaḥa* to be salty), the name of the site of Ḥims, a district into which the Jewish community of Fez were compelled to move in 1438. Cf. French *mellah* (1860).]

Completely new etymological information has been provided for the many new entries of exotic origin. While this concentrates on the transmission of the loan into English rather than the further etymology of the word in the source language, there are examples here of etymologies that have never before been published in any language.   *Joanna Tulloch*

References

The Century Dictionary (1889-91), ed. by William Dwight Whitney. London and New York.

The Dictionary of National Biography (1885-1901), ed. by Sir Leslie Stephen and Sir Sidney Lee. London.

Durkin, Philip (1999). 'Root and branch: revising the etymological component of the OED'. *Transactions of the Philological Society*, 97: 1-50.

Gmelin, Leopold (1844). *Handbuch der Chemie* (4th revised edn.). Heidelberg.

— (1848-72). *Hand-book of chemistry by Leopold Gmelin* (translated by Henry Watts). London.

Littré, Émile (1863-72). *Dictionnaire de la langue française* (30 fascicles = 1881, 5 vols.). Paris.

Nicolls, Thomas (1550). *The Hystory writtone by Thucidides the Athenyan of the warre which was betweene the Peloponesians and the Athenyans translated out of Frenche into the Englysh language by Thomas Nicolls citizeine and Goldsmith of London.* London.

Nietzsche, Friedrich Wilhelm (1888, 2nd edn. 1892). *Der Fall Wagner. Ein Musikanten-Problem.* Leipzig.

— (1896). *The Case of Wagner. - Nietzsche contra Wagner. - The Twilight of the Idols. - The Antichrist* (translated by T. Common: *The Works of Friedrich Nietzsche*, vol. 11). London.

The Oxford English Dictionary (1884-1933), ed. by Sir James A. H. Murray, Henry Bradley, Sir William A. Craigie and Charles T. Onions; *Supplement* (1972-86), ed. by Robert W. Burchfield; 2nd edn. (1989), ed. by John A. Simpson and Edmund S. C. Weiner; *Additions Series* (1993-7), ed. by John A. Simpson, Edmund S. C. Weiner and Michael Proffitt; 3rd edn. (in progress) *OED Online* (March 2000-), ed. by John A. Simpson. Oxford.

Seyssel, Claude de (1545). *L'histoire de Thucydide Athenien de la guerre qui fut entre les Peloponesiens & Atheniens.* Paris.

Trésor de la langue française: Dictionnaire de la langue du XIXe et XXe siècle (1789-1960) (1971-94), ed. Paul Imbs and Bernard Quemada (16 vols.). Paris.

Webster's New Intenational Dictionary of the English Language (1911 edn.), ed. by W. T. Harries and F. Sturges Allen. London and Springfield, Mass.

Representations of Indo-European lexical dissimilarities

Tanmanjeet Singh Dhesi and Mario Cortina Borja[*]

1. Introduction

The classification of languages within a family using lexical evidence has a long tradition, going back even farther than Sir William Jones' famous remark, in the 18th century, on the affinities between Sanskrit, Latin, and Greek. There are two broad methods for classifying languages, classical and statistical. Classical techniques, employed in philology, make use of phonological and morphological comparisons to infer groupings of languages. Statistical techniques, pioneered by Swadesh (1972), analyse the lexical cognate matrix in order to elucidate the relations of proximity contained in it. We do not follow Swadesh's approach of glottochronology. The statistical methods presented in this paper can be divided in two classes: tree-based methods and geometric methods. They both explore the neighbourhood relations contained in the cognate matrix; their main differences are in their output. In the former it will be a tree showing the different branches of the family according to a particular definition of distance between groups of languages, as in classical cluster analysis methods (Anderberg 1973), or a phylogenetic tree, obtained by minimizing a functional (e.g. sum of squares) defined in terms of geometrical inequalities (Everitt and Rabe-Hesketh 1997). Geometric methods yield a configuration, typically in two or three dimensions, in which the points represent the languages and the distance relations shown there mirror those present in the lexical dissimilarities (Cox and Cox 1994). In this paper we will focus on the usage of several methods of statistical classification to obtain representations of lexical data. We ponder their advantages and disadvantages, and illustrate them using the lexical data from the Indo-European family prepared by Dyen *et al.* (1992). Their work is our starting point.

The main purpose of the figures in this paper is to display the groups' positions rather than those of individual languages - relations among languages are described in the text. In figure 3 the languages' names have been omitted. In figures 2, 4, 5, 8, and 9 it would be difficult to improve the clarity of the 95 labels and this has not been attempted. In figures 6 and 7 this would have been impossible as the relative positions of many languages are too close to each other.

2. Data

When constructing a classification of languages within a family, one does not set-out with a list of features that related languages should satisfy if they were indeed related. In fact, there should be no a priori restrictions. However, we would require a reasonably long list of systematic commonalities (not individual forms), for which the kind of commonality required is not pre-determined. This is provided by the Swadesh list of cognates. Two forms are cognate if they have descended in unbroken lines from the same 'ancestor'. To establish cognation, one has to demonstrate a common ancestor for a word in several languages based on a system of changes. The Swadesh list corresponds to words that represent a basic vocabulary; in this paper we used the 200-words list for 95 languages from the Indo-European phylum and their corresponding matrix of lexicostatistical percentages for all 4465 pairs of languages. These figures were calculated removing the discoverable loanwords. Professor Dyen and her collaborators have kindly made the data available through the internet.

Comparative lexicostatistics refers to methods (in particular language trees) for the investigation and hypotheses' development of historical relationships between languages, via the use of cognates. It generally uses the lexicostatistical percentage, i.e. the proportion of basic meanings for which words in two given languages are cognate. Based on the Swadesh 200-words list, it can be more precisely described as the 'percentage of homosemantic cognation among the basic vocabulary'. It estimates the nearness relationship between two languages — a high percentage figure indicates a fairly recent divergence from a common ancestor, i.e. a closer relationship.

A criticism levelled at lexicostatistics is that it is unrealistic to assume that different meanings have the same replacement rate, i.e. that any two words used for different meanings are equally likely to change. Nonetheless, the approach does give a good approximation to classifications obtained using classical methods. Improvements for a more realistic approach can be attained, but there is a trade-off between realism and additional complexity (not to mention computational problems), as discussed by Guy (1980) and Embleton (1986).

Our main aim is to explore as many interesting representations of the data as possible. Linguists, using the historical linguistic construction method, have always spoken of families of languages, each with certain subfamilies. The methods used in this paper should be a powerful tool to help the linguist to ascertain the position of groups of languages within a phylum.

Traditional classification in the Indo-European phylum, based upon historical evidence and linguistics, has identified the following groups: Celtic (7 languages of those analysed in this paper), Romance (16), Germanic (15), Balto-Slavic

(26), Albanian (6), Greek (5), Armenian (2), Indic (11), and Iranian (7) as the major branches of the family. The data analysed in this paper consist of 19,000 records; that is to say 200 cognates (based on the Swadesh 200-list) for each of the 95 speech varieties from the Indo-European family listed in Table 1 in the order given by Dyen *et al.* (1992). Appendix 1, also taken from Dyen *et al.* (1992), contains the full names of the languages and their sources. Throughout the paper we use the names listed in Table 1.

Table I: The 95 Indo-European languages analysed

IrishA	Spanish	LithuanianO	Marathi	Tadzik
IrishB	PortugueseST	LithuanianST	Gujarati	Baluchi
WelshN	Brazilian	Latvian	PanjabiST	Wakhi
WelshC	Catalan	Slovenian	Lahnda	AlbanianT
BretonList	GermanST	LusatianL	Hindi	AlbanianTop
BretonSE	PennDutch	LusatianU	Bengali	AlbanianG
BretonST	DutchList	Czech	NepaliList	AlbanianK
RumanianList	Afrikaans	Slovak	Khaskura	AlbanianC
Vlach	Flemish	CzechE	GreekML	RUSSIANP
Italian	Frisian	Ukrainian	GreekMD	UKRAINIANP
Ladin	SwedishUp	Byelorussian	GreekMod	BYELORUSSIANP
Provencal	SwedishVL	Polish	GreekD	POLISHP
French	SwedishList	Russian	GreekK	SLOVAKP
Walloon	Danish	Macedonian	ArmenianMod	CZECHP
FrenchCreoleC	Riksmal	Bulgarian	ArmenianList	SLOVENIANP
FrenchCreoleD	IcelandicST	Serbocroatian	Ossetic	SERBOCROATIANP
SardinianN	Faroese	GypsyGk	Afghan	MACEDONIANP
SardinianL	EnglishST	Singhalese	Waziri	BULGARIANP
SardinianC	Takitaki	Kashmiri	PersianList	ALBANIAN

Thus the dataset is a 95 x 95 lexical dissimilarity matrix. Note, however, that 200 meanings (for cognates) were not available for all the languages. Nevertheless, this should not have made a significant difference, since at least 190 meanings were available for all but four of the languages, with the lowest being 183 (for Vlach and GypsyGk).

3. Methods

3.1. Tree-based Methods

Tree-based methods produce a more structured illustration of the data objects than other techniques, including geometric methods and for this reason alone, tree representations are likely to be preferred by most linguists. Trees can be used to produce natural and insightful models for the evolution of linguistic diversification; however in this paper we are only concerned with an approximation to the outcome of these complex processes. Ross (1997) and Lohr (2000) discuss family tree models in contexts similar to ours.

3.1.1. Hierarchical and divisive algorithms

Hierarchical algorithms combine or divide existing groups to produce a tree displaying the order in which groups are merged or divided. For agglomerative nesting, two methods defined by Kaufman and Rousseeuw (1990) are of particular note: `hclust` (shorthand for hierarchical clustering) and `agnes` (id. for agglomerative nesting). The first method joins those two clusters with the smallest between-cluster dissimilarity and applies this rule until there is only one cluster. There are several ways to calculate the dissimilarity between clusters; the basic forms are: single linkage (by taking the nearest neighbour between pairs of points from different clusters), complete linkage (by taking the furthest neighbour), and average linkage, which obtains the mean of all the possible pairs from different clusters. The last form is supposedly more robust and consistent. Other forms consider a weighted average or the median of the between-cluster dissimilarities. The `agnes` procedure is particularly good, since it calculates an agglomerative coefficient, which measures the amount of clustering structure in parts of the dataset. In addition, it produces a banner plot, which shows the structure of the successive mergers.

Another example of a tree-based method is Kaufman and Rousseeuw's `diana` (for divisive analysis) which starts at the opposite end of `hclust` and `agnes`, with one group, and divides it until each object is a different cluster. A banner plot is also produced, along with a divisive coefficient, which measures the clustering structure of the data.

Partitioning clustering algorithms include the very useful Partitioning Around Medoids (`pam`). This method starts with a fixed number of clusters and then chooses the group identity of a particular subject by identifying its nearest medoid, which is a centrotype for the cluster. This implies that the optimization is done with respect to k objects (Kaufman and Rousseeuw 1990), and `pam` therefore can produce an optimal partition of the objects into k groups, as well as

yielding the medoid objects for each group. In addition, the user is presented with an index that can be used to determine the optimal number of clusters, which varies from 1 to *n* (the number of objects). A silhouette plot is also produced, showing the partitioning of the data. An average silhouette width is given, with the values ranging from -1 to 1. A silhouette width of 0 for any object indicates that it lies between two clusters, -1 implies a very badly classified object, and 1 indicates a very well classified object. This plot is a useful tool to determine the best-defined clusters in the dataset.

3.1.2. Phylogenetic trees

Clustering techniques are based on agglomerative and divisive algorithms. The additive and ultrametric phylogenetic trees, meanwhile, are based on minimizing a given objective function, usually a sum of squares-type criterion, subject to some conditions, which usually refer to sets of inequalities defined amongst subsets of dissimilarities.

In the notation of Everitt and Rabe-Hesketh (1997), an additive tree is a connected, undirected graph where all pairs of nodes are connected by a unique path. Its defining feature is a relationship between the path length distances, known as the additive inequality, or the four-point condition. Any four nodes *a,b,c,d* in a set must satisfy the following inequality for a set of distances to define an additive tree:

$$d_{ij} + d_{kl} \leq \max (d_{ik} + d_{jl}, d_{il} + d_{jk}) \text{ for all } i, j, k, l$$

Again, following the notation of Everitt and Rabe-Hesketh (1997), an ultrametric tree is an additive tree in which a terminal node is equidistant from some specific node, called the root of the tree. All triples in a set of observed dissimilarities must satisfy the following ultrametric inequality to be represented by an ultrametric tree model:

$$d_{ij} \leq \max (d_{ik}, d_{jk}) \text{ for all } i, j, k$$

Examples of an additive tree and an ultrametric tree appear in Figure 1. We took advantage of the set of programmes within the package `phylip` (Phylogeny Inference Package), written by Professor Joe Felsenstein (1993), who has kindly made it available through the internet. In particular, we used the programmes `fitch` to fit additive trees, and `kitsch` for ultrametric trees to the lexical dissimilarities. The `drawgram` and `drawtree` executables within `phylip` actually draw the rooted and unrooted trees respectively. The two executables are quite powerful, since one can look at trees from different angles and choose

different types of diagrams (such as a `swoopogram`, `curvogram`, and `eurogram`), as well as a whole host of other options.

Figure 1: Examples of additive and ultrametric trees
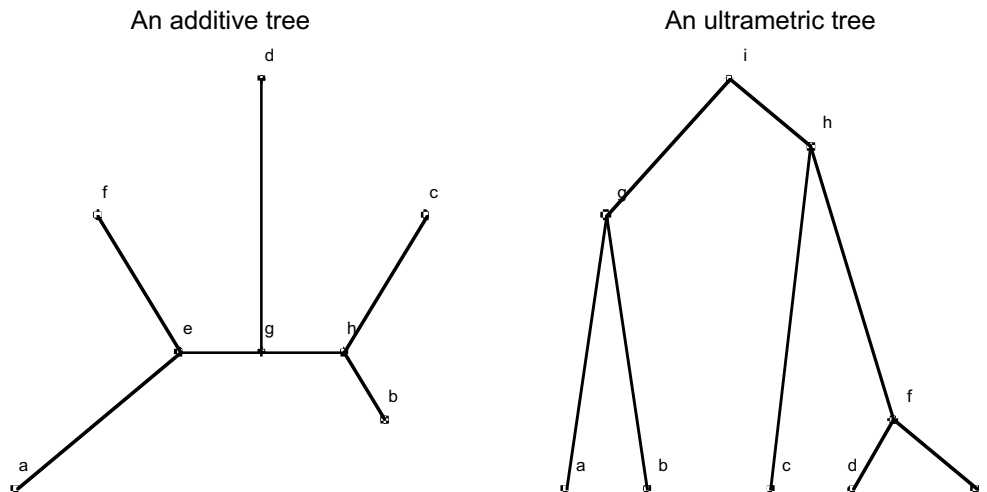
3.2. Geometric methods

Most geometric methods can be referred to as distance methods. These produce a configuration of points, representing in most cases a 'map' of the data objects. The nearness relationships amongst the objects in this fitted map are distances, not dissimilarities, i.e. they satisfy the triangle inequality:

$$d_{ij} \le d_{ik} + d_{jk} \text{ for all } i, j, k$$

Note that any distance is a dissimilarity, though the converse is not true. Distance-based methods are often known as metric methods.

Given a dataset of *n* observations over *p* variables, a visual aid would be *n* points in *p*-dimensional space. Multi-dimensional scaling (MDS) attempts to find a configuration in *d*-dimensions, where *d* is less than *p*. Typically, *d* = 2 or *d* = 3. Before performing MDS, the data have to be transformed into a dissimilarity matrix. Distances between points in the lower dimensional configuration should match those of the original dissimilarity matrix, in the sense that proximity information is preserved as much as possible.

A measure for the goodness of fit of the representation is its stress value, which needs to be minimized (see Venables and Ripley 1999). This is a suitable criterion, since it has certain desirable properties, such as invariance under translation, and being an orthogonal transformation. An example of a general

criterion is the one proposed by Kruskal and Shepard in the 1960s, which constructs configurations minimizing

$$STRESS^2 = \frac{\Sigma_{i \neq j}\left(\theta\left(d_{ij}\right)-d^*_{ij}\right)^2}{\Sigma_{i \neq j}\left(d^*_{ij}\right)^2}$$

where d_{ij} is the original dissimilarity between objects i,j, d^*_{ij} is the fitted distance in the lower dimensional configuration, and θ is an increasing function.

Classical MDS uses a Euclidean distance, which assumes some metric structure that the data may not possess. A non-metric alternative is Sammon's (1969) non-linear mapping which minimizes the weighted stress function:

$$STRESS^2 = \frac{\Sigma_{i \neq j}\left(d_{ij}-d^*_{ij}\right)^2 \big/ d_{ij}}{\Sigma_{i \neq j}\left(d_{ij}\right)}$$

We used Venables and Ripley's (1999) iterative algorithm to fit configurations with Sammon's method.

Classical/metric scaling, known as principal coordinate analysis, preserves to a great extent the distances between the observations. In contrast, non-metric scaling, such as Sammon's non-linear mapping, is based upon monotonic regression, which preserves only the rankings of the data. Non-metric scaling is used when there is subjectivity present in the data. It could be argued that the compilation of our dataset most probably involved some subjective evidence and hence the Sammon mapping and Kruskal's non-metric MDS come into play. For these techniques, the plots may appear in a planar 2-D map, but the surface must be thought of as being non-linear.

Another non-metric geometric method is Friedman-Rafsky's (1981) 'planing' algorithm which is based on the minimum spanning tree. A minimum spanning tree is a series of edges which join all points to form a connected graph of minimum total length. The local preservation of distances is the key feature for this plot. It is not as accurate as Sammon, but is much faster usually; however it can only be used to produce 2-dimensional configurations. Note that this method does not minimize a stress criterion, but looks for 2-D configurations such that their Euclidean distances locally resemble the original dissimilarities.

In the next section we present examples of these methods applied to the Indo-European lexical dissimilarities matrix.

4. Analysis of Indo-European lexical dissimilarities

Several techniques were attempted for the Indo-European phylum. However, not all of them are listed; only those producing representations deemed to have a good trade-off between internal homogeneity and within-groups heterogeneity have been shown. The decision of whether or not any representation should have been shown depended on quantities particular to the corresponding method, e.g. stress, agglomerative coefficient, proportion of explained variance as a function of dimensionality, silhouette width, proportion of additive and ultrametric inequalities satisfied by the reconstructed configuration, etc. as well as on qualitative judgements, e.g. absence of chained groups.

4.1. Tree methods

Classification is not an exact science and it is possible to have languages which obviously belong to different families appearing in the same cluster. The best illustration out of all the hierarchical and divisive clustering techniques attempted was the clustering tree obtained via the use of function `hclust`, as shown in Figure 4.1. Recall that the group average method for calculating between-cluster dissimilarities is usually more robust and consistent. Applying `agnes`, an agglomerative coefficient of 0.79 was obtained, confirming a good clustering. Though other clustering methods (e.g. `diana`) yielded similar values of this coefficient, we chose the one produced with `agnes` as it gave the clearest visual representation. After a re-ordering of the clustering tree to display the clusters more clearly, one sees that languages have been allocated to each branch, exactly according to traditional classification. The fact that the Albanian speech varieties are very dissimilar to all the others is again emphasized by their branch being merged last in the tree. We may speculate that this happened as a result of having their inherited cognates replaced by Latin loanwords relatively recently. The tree in Figure 2 indicates that our agglomerative methods do not replicate exactly the historical diversification processes in a straightforward way.

If a height of 0.76 is taken, nine branches are obtained (if we ignore that Ossetic is very different amongst the Iranian group). These are precisely those nine identified by traditional linguists as being the principal groups in the phylum. If we were to separate Balto-Slavic into two branches we would have had to split the tree at a height which would imply separating two Celtic branches (Irish and non-Irish), as well as recognizing some of the Indic and Iranian languages as single branches. Note the position of English, which is equidistant from the Dutch-German and Nordic group. This agrees with the results obtained by Dyen *et al.* (1992) and implies that English cannot be

grouped with Frisian if one uses modern data. Dyen *et al.* (1992) also stated the 'failure to relate English immediately with Frisian in accordance with the Ingveonic hypothesis'. Our analyses thus reveal extended contact as much as initial speciation, and suggest both a process of lexical diffusion from Dutch and Low German into Frisian, as well as an effect of the isolation of English with respect to the continental West Germanic languages. It is also interesting how the Balto-Slavic group is closer to the Germanic and Romance branches than Celtic, which is geographically closer. This suggests again a process of extended contact. Ross (1997) discusses in detail contact-induced change processes and models to describe innovation-defined subgroups within a phylum.

Figure 2: Hierarchical clustering

Agglomerative hierarchical clustering (hclust): group average
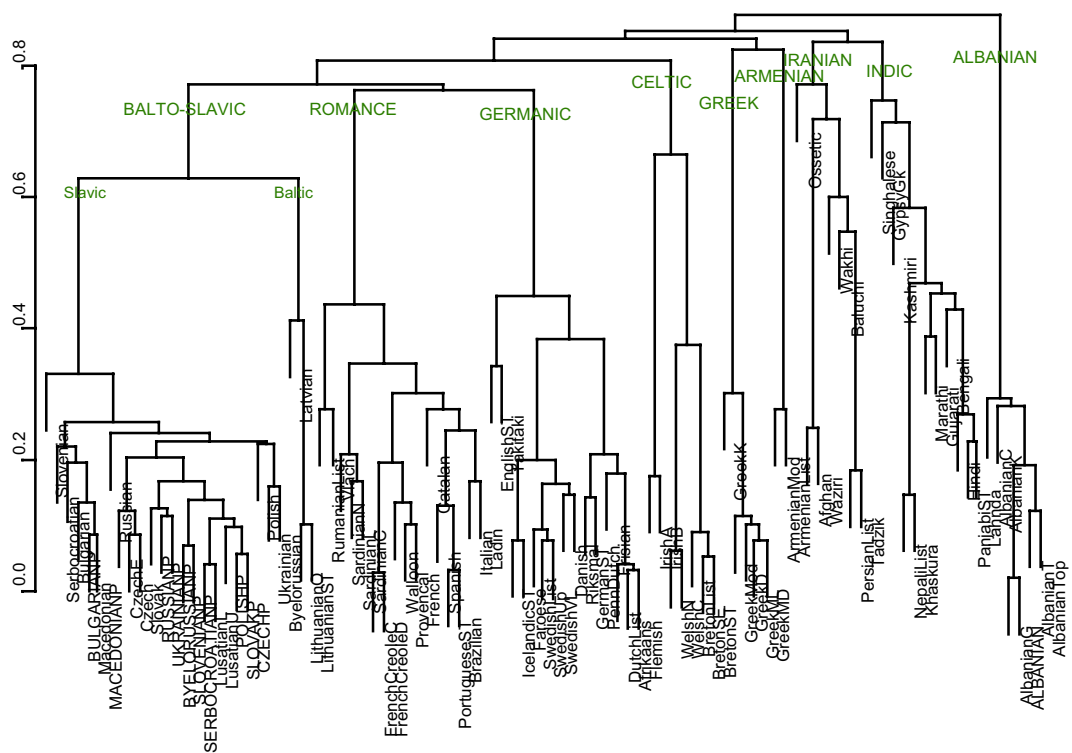
With 0.82 as our height, five cohorts are present. One of these is the combination of Balto-Slavic, Romance, Germanic and Celtic forming one group. This would add force to the argument of Meillet for a north-western group of Indo-European languages. A height of 0.85 would imply the subdividing of the family into three major groups: Albanian, Indo-Iranian, and the rest making the third group. Hence, contradicting Dyen *et al.* (1992), one could say that there is evidence of an Indo-Iranian group.

After implementing `pam` several times for different values of *k* (the number of clusters specified), the partitioned clustering shown in Figure 3 was attained. The language names only serve to clutter the representation and hence have not been included. Instead, the name of each branch has been displayed. The ellipses from method `pam` are similar to the confidence intervals for each cluster. This is probably why the Balto-Slavic ellipse is the largest in size; it indicates the inherent subdivision of the Baltic and Slavic, which is also visible on the plot. It is misleading to emphasize, or even attempt to interpret the two components plotted, which explain just 32.29 of the total variability. The important feature of this representation is the clustering, which is very informative. If the `pam` graph is to be believed the linguistic classification can be regarded as perfect, since all the speech varieties have been classified exactly according to it.

Figure 3: Clustering with Partitioning Around Medoids method
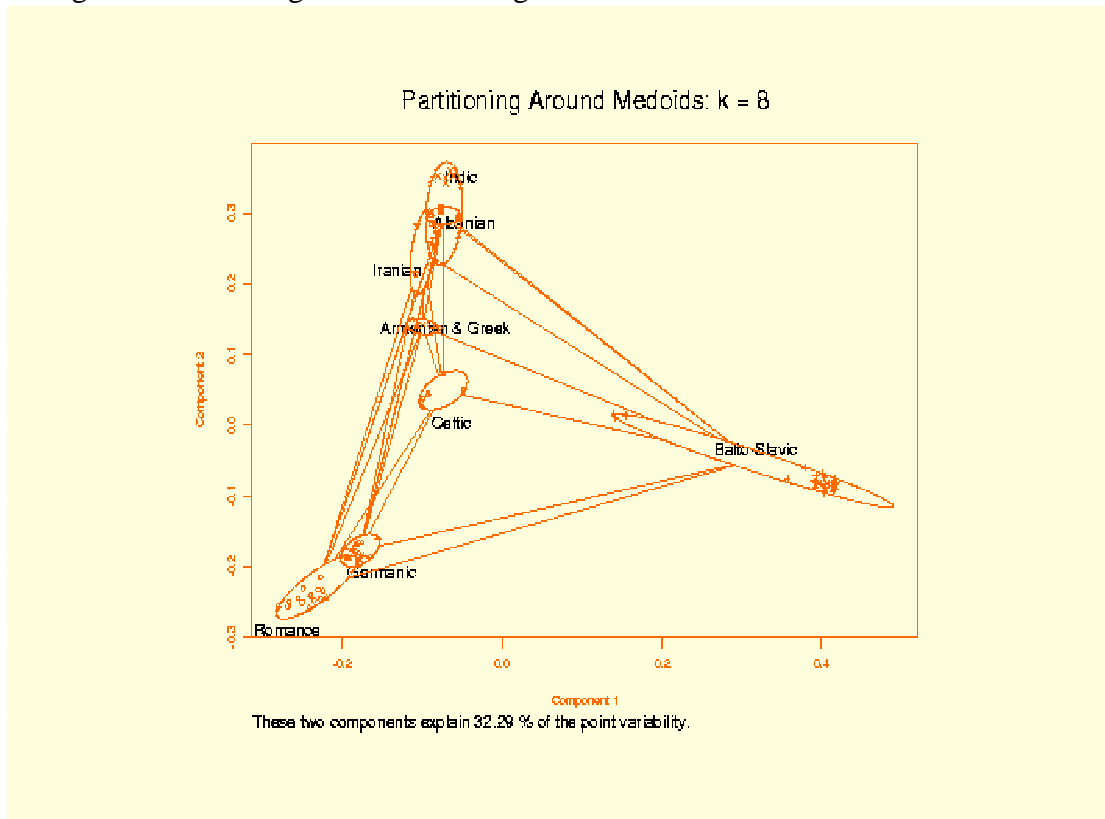
An average silhouette width of 0.51 was obtained, which indicates quite a good clustering. Recall that the silhouette width for an object, ranging from -1 to 1, shows how well classified the object is within the cluster, where -1 indicates a bad classification and 0 suggests that the object lies between two clusters. (So, we would like the average silhouette width to be as large as possible.) The highest average silhouette width of 0.55 was obtained for *k* = 13, but the clustering produced too many groups with very few languages in them. We

therefore applied a trade-off argument between minimal within-cluster variance (i.e. having internally consistent groups) and maximal between groups variance (i.e. having well-defined groups). The average width increased until $k = 8$ and then decreased slightly, before reaching a peak at $k = 13$. In this case, it is reasonable to have the best clustering defined by the first peak. It also agrees with the hierarchical clustering results described previously.

The Iranian languages had the highest average dissimilarity of any cluster, with a figure of 0.48, which again corresponds with the results of the hierarchical clustering. Recall that Ossetic was very different from the other six Iranian languages, while out of the remaining six, Afghan and Waziri were again a distinct pairing, distant from the others. In contrast, as one would expect, the Albanian languages formed a tightly-knit group, with an average dissimilarity of 0.16.

In order to compute how well the Indo-European lexical distance matrix satisfied the additive and ultrametric inequalities we examined all possible subsets of four and three languages respectively. The number of additive inequalities satisfied for the data were 2,544,550 (79.9%). The corresponding figure for ultrametric inequalities satisfied was 101,920 (73.64%). This was reflected in the quality of the ultrametric trees; the unrooted ultrametric tree, for example, had the Armenian, Greek and Iranian branches intertwined. Hence, the additive trees were preferred.

The rooted additive curvogram in Figure 4 shows the tree constructed with respect to IrishA. Any language could have been chosen as the reference point. Taking IrishA produced a visually clear representation perhaps because its distances with the rest of the languages have a relatively small variance: i.e. it is a 'central' language, in the sense that it is approximately uniformly near to the rest of the languages. Cortina Borja and Valiñas Coalla (1989) have discussed the idea of centrality based on variation of lexical distances within a phylum. As expected, IrishB is the closest language to IrishA, followed by the other Celtic languages in one group. Again, the representation distinguishes between the Breton and Welsh varieties. All the remaining branches are clearly identifiable and there is very little difference in their distance from IrishA, though the Indo-Iranian and Albanian cohorts are probably a little more distant. This assessment is in perfect agreement with the features identified for rank-distance graphs (not included here). In essence, the rooted curvogram operates along similar lines to the rank-distance graphs, but is a more elegant technique which produces clearer representations. It is also more flexible as the data set can be viewed from the point of view of any language.

Figure 4: Rooted additive curvogram
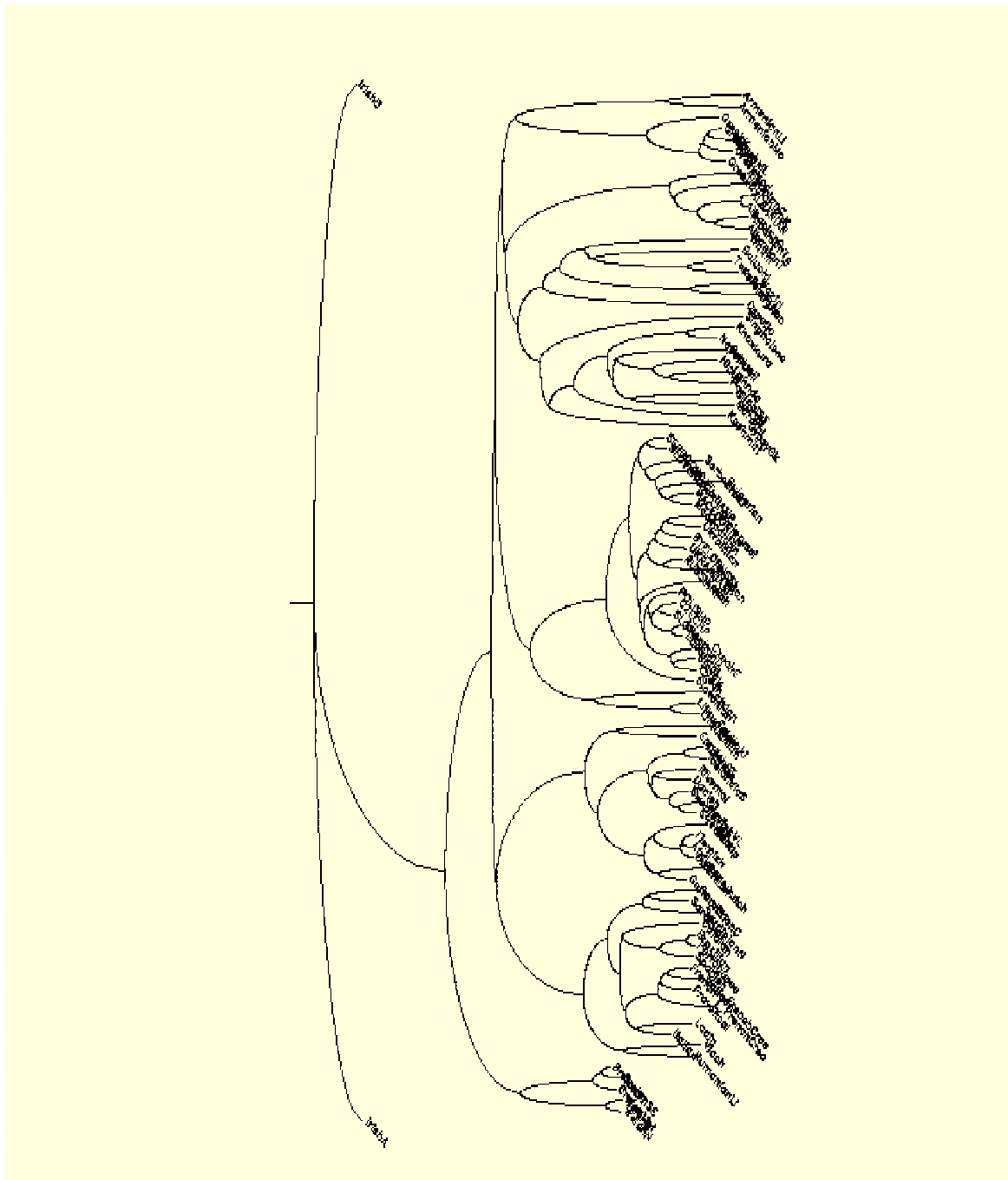
The unrooted additive tree in Figure 5 concurs perfectly with the previous representations, especially the hierarchical clustering tree. Groups are once more easily distinguishable. However, the really outstanding feature of the representation is the degree to which the details of this tree match the previous clustering tree.

Figure 5: Unrooted additive tree
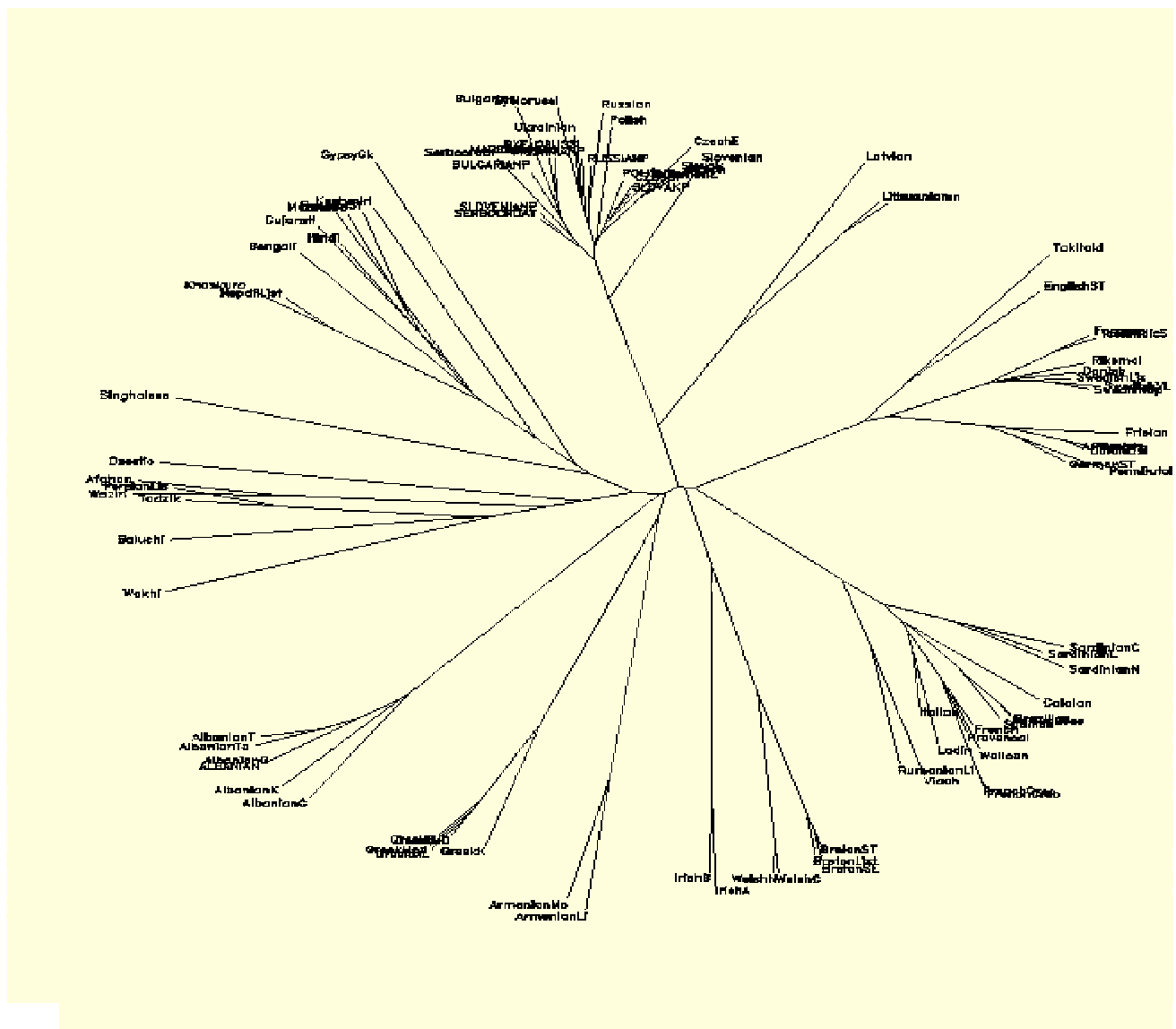
Ossetic is quite distinct amongst the Iranian languages, followed by the Afghan and Waziri, which are at a fair distance from the rest. Likewise, amongst the Indic languages, GypsyGk is the most different. The compactness of the Albanian group is all too apparent, while it is also shown that Greek and Armenian may be the closest to each other, but this still represents a large distance. The proximity of English and its pidgin Takitaki (a language from New Guinea) is also illustrated.

4.2. Geometric methods

Metric scaling in 2-D produced the most informative results, with respect to all the other metric MDS plots (Figure 6). Seven of the easily identifiable groups have been labelled in the configuration shown. The most curious aspect about it is the alignment along the first component, i.e. x-axis. Four x-coordinate values

(treating the Romance and Germanic as generating one value of approximately 0.25) can be seen, but it is difficult to interpret its meaning. The metric scaling had a stress value of 0.37, indicating that the goodness-of-fit is not as good as we would like.

Figure 6: Metric multi-dimensional scaling

Indo-European family: metric scaling
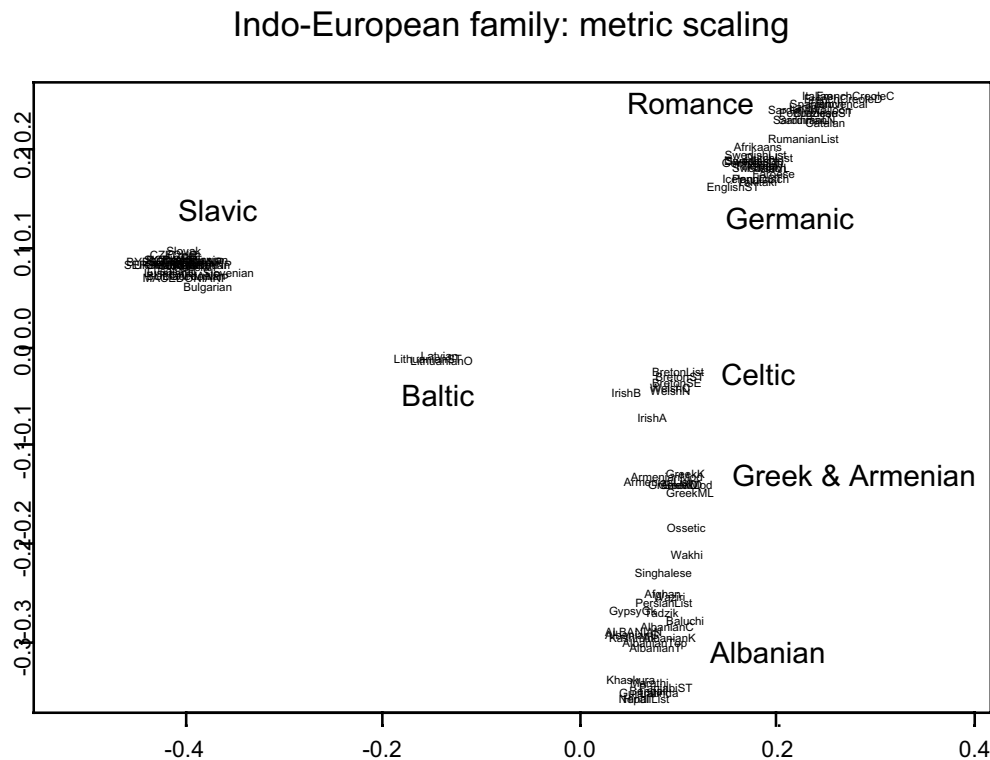
The Albanian languages were interspersed with the Indo-Iranian languages. A 3-D MDS plot using the `spin` option in S-PLUS led to five fairly distinct clusters, within which the Greek/Armenian, Albanian, and Indo-Iranian groups were not clearly identifiable. The global/phylum structure, which may be obvious, could well have been suppressing the local, i.e. branch structure. Hence, a decision was taken to perform MDS on each of these subgroups. A plot for the Gaelic sub-family showed a polarized illustration with three groups (Breton, Irish, and Welsh) in three different corners. The 'West-European' configuration presented the Romance languages in a tight group on the left, while the two Dutch-German and Nordic groups were separated on the right. The most interesting representation, however, was for the 'South-East Europe and Indo-Iranian' grouping, as shown in Figure 7. The European-Asian divide is evident, and five distinct cohorts are noticeable. Unlike the full metric-scaling graph (Figure 6), the x-axis signifies longitude. From the European Albanian and Greek groups,

one can proceed in an easterly direction across Armenia, Iran and then into India.

Figure 7: Metric MDS for South East Europe and Indo-Iranian grouping

South-East Europe and Indo-Iranian grouping: metric scaling
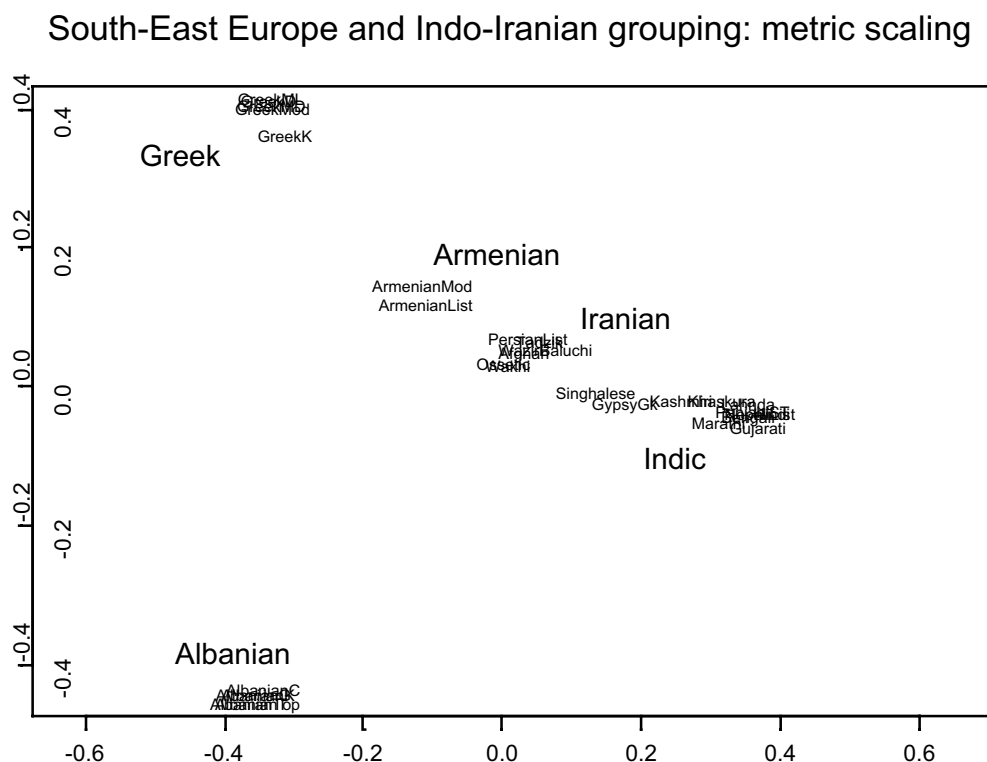
An unambiguous Asian and European distinction is presented by the Sammon mapping in Figure 8 with the exception of Albanian, which once again poses problems for a clear-cut classification. The representation is the best 2-D configuration (in terms of trade-off arguments) amongst all the distance methods (including Kruskal's MDS and Friedman-Rafsky's minimum spanning tree). The European and Asian families can be separated in terms of higher dimensional 'horseshoes', as defined by Kendall (1971), which indicate the non-linear nature of the true configurations; each of the cohorts is clearly visible. Only Ossetic, with the Armenian varieties, seems to be outside its prescribed group. Ossetic has been noted as being quite distinct amongst the Iranian languages, but this is the first time that it is so strikingly detached from its group. The Armenian and Celtic groups, meanwhile, may seem close in Figure 8, but it should be taken into account that Sammon representations are non-linear, in the sense that the surface on which the points lie is not planar. The stress value for the 2-D Sammon mapping is 0.103, which indicates an excellent

fit, though too much cannot be read into this since the interpretation of stress is strongly dependent on the number of stimuli involved (Everitt and Rabe-Hesketh 1997).

Figure 8: Sammon's non-linear mapping

Indo-European family: Sammon mapping
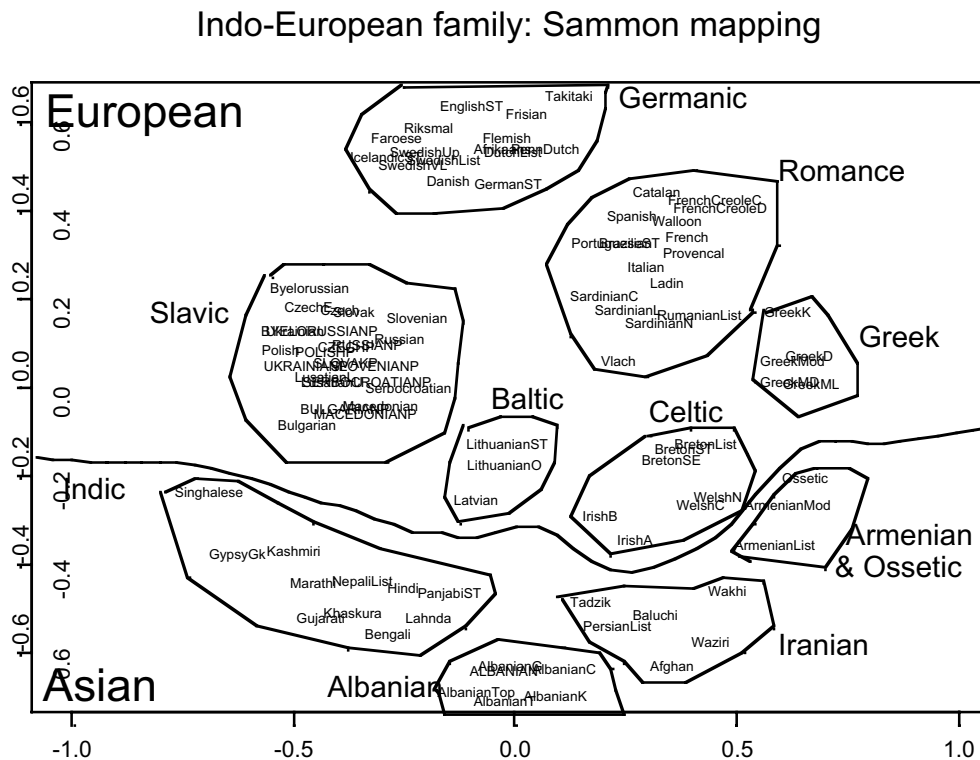
Amongst the 3-D representations constructed, the one with best trade-off characteristics was the isotonic MDS shown in Figure 9.

Recall that the isotonic regression fitting preserves the order and not the values, in contrast to metric scaling, which tries to fit the values themselves. Vertical lines were constructed, so that the reader could get a feel for the 3-D positioning of the groups. With regards to the stress values, 3-D isotonic MDS had a stress of 11.49, which indicates a fair fit. As the number of dimensions were increased, the stress values decreased, with 6-D isotonic MDS having a stress value of 4.61, though this is obviously of no use to us in our search for a representation.

Overall, for the Indo-European phylum, the geometric analysis corresponded well with the tree techniques. All the representations were very good, but the hierarchical clustering and unrooted tree were arguably the most informative about the group dynamics.

Figure 9: Isotonic multidimensional scaling

Indo-European family: 3-D Isotonic MDS
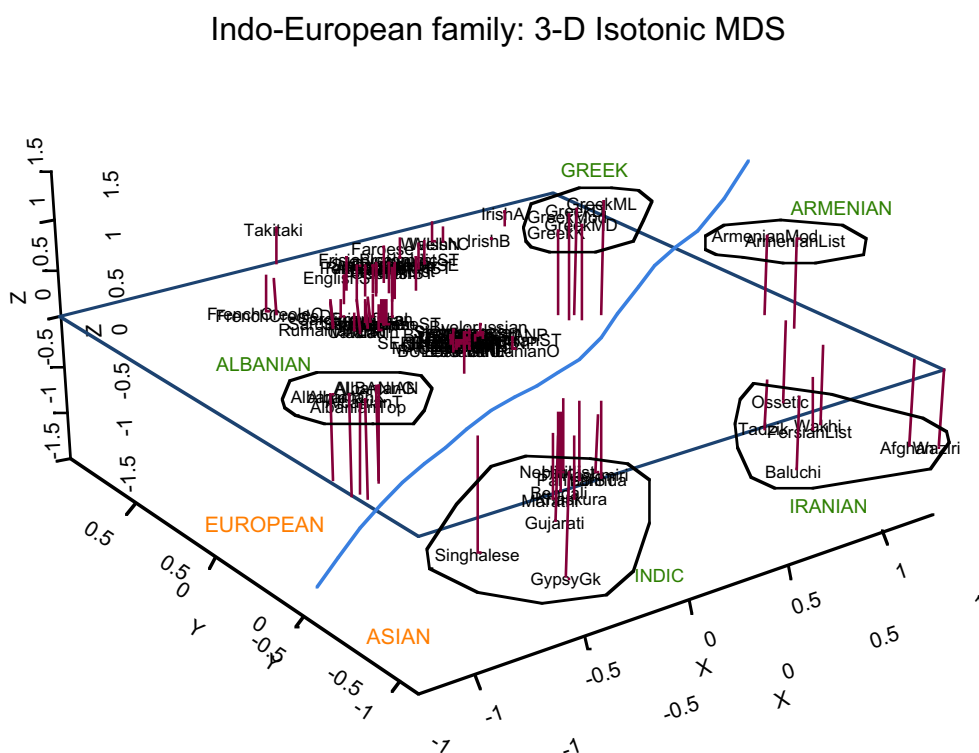
5. Discussion

Lexicostatistics was applied to numerous Indo-European languages. Comparisons were then made with the historical linguistic construction method (the traditional technique). In general, the lexicostatistical classifications were consistent with the widely acknowledged classifications. This correspondence reinforces the general validity of the lexicostatistical method. The advantages of using a series of analytical methods to explore the structure of lexical data were shown.

Tree-based methods were in almost perfect agreement within themselves, as well as the traditional classification. The hierarchical clustering and unrooted tree were arguably the most outstanding. One point to make is that, using a battery of methods on modern data, we were able to recognize the existence of an Indo-Iranian clade (from the cluster analysis), something which Dyen *et al.* (1992) did not achieve. Historical evidence seems to point to the existence of such a group, and our result shows the importance of using a wide range of methods in analysing historical lexical data.

Sammon's non-linear mapping produced the better configurations for two dimensions, while the metric scaling MDS was better in three dimensions.

Phylogenetic trees have been used recently by Lohr (2000) to analyse lexical and phonological data from a set of eleven Indo-European languages. For both procedures it is important of course to check the goodness of fit of the configurations or trees fitted using a stress measure and the proportion of additive and ultrametric inequalities satisfied, respectively.

There were certain outliers amongst the languages analysed. The Albanian and Armenian groups spring to mind. The methods discussed in this paper can only point out their position relatively to the rest of the phylum.

The study of languages and how they have spread is a complex subject. It does not involve traditional linguistics alone. It has been argued (Renfrew 1992, 1998) that there appears to be a remarkable potential for the synthesis between pure linguistics, archaeology, and modern genetics. Ancillary information can be obtained from archaeology, though there is debate about which prehistoric records can be used to illustrate the social and economic processes underlying linguistic change. The significance of agriculture deserves special mention given the link between language replacement and agricultural dispersal. Language replacement usually arises due to the more advanced technology of the incoming group, who may be small in number, but are able to exert their dominance by the use of force and/or technology. The concept of 'cultivators expanding and absorbing foragers' is taken for granted, since a stable farming community is associated with expansion.

The proposal of a possible relationship between genetics and linguistics dates back to the time of Darwin, who saw the possible equivalence between human descent and language evolution. While one would assume that correlations between language and gene frequency exist, the issue is rather convoluted by the presence of other factors. Language replacement, using the elite dominance model (Renfrew 1992), where the incomers replace the existing elite by the use of arms, is a fine example. The incomer group may be able to change the language, but in most cases they cannot significantly alter the genetic make-up of the indigenous people, owing to their small number. Yet another complicating factor is the fact that in such cases most incomers are male, which carries with it many genetic complications. According to Cavalli-Sforza *et al.* (1994), it is also possible that extensive gene replacement has transpired via the prolonged contact and gene flow from neighbours, without any change in language (see also Piazza *et al.* 1995). It would be fascinating to conduct a study on genetic and lexical data. A more ambitious enhancement, following Embleton's (1986: 169-70) multi-faceted approach to linguistic reconstructions, would analyse data including many components/variables, such as lexical, phonological, morphological, grammatical, genetic, archaeological, and geographical information.

Appendix: the 95 Indo-European languages analysed and their sources

The appendix has been taken literally from Dyen *et al.* (1992). The listing is in alphabetical order.

The Swadesh meanings for the ten Slavic languages written in capital letters ending with a P in Table 1 were obtained from Fodor (1961).

Finally, Dyen *et al.* (1992) point out that the last list in Table 1, denoted ALBANIAN, was virtually identical with another Albanian list, except for having some missing forms. (These two lists were from supposedly different Albanian dialects.) They do not specify the source of this list nor which of the other Albanian languages it is closest to. For completeness's sake we included all the data presented in the monograph by Dyen *et al.*

The number of Swadesh meanings glossed in each list is shown in parentheses after the name of the list.

Afghan (200): P. B. Zudin, *Russko-Afganskij slovar'*. Moscow, 1955.

Afrikaans (200): T. Kriel, *Die Nuwe Afrikaanse Skool-Wordenboek*. Kaapstad, 1945.

AlbanianC (190): Piana degli Albanesi, Sicily; Eric Hamp.

AlbanianK (200): Sophikon, Korinthia, Greece; Eric Hamp.

AlbanianG (190): Geg dialect; G. Weigand, *Albanesisch-deutsches, deutsch-albanesisches Wörterbuch*. Leipzig, 1914.

AlbanianT (198): Tosk dialect; Georges Schmidt.

AlbanianTop (200): Ogrén/Përmét; Eric Hamp.

ArmenianList (190): Adapzar dialect; V. Gerrard.

ArmenianMod (196): Eastern Modern Armenian; Ilya Abuladze, in Knut Bergland and Hans Vogt, 'On the validity of glottochronology', *Current Anthropology*, 3.2 (1962): 115-53.

Baluchi (193): G.W. Gilbertson, *English-Balochi Colloquial Dictionary*. Hertford, 1925.

Bengali (198): Frank Southwork.

Brazilian (197): M. Reno, *Portuguese: A Handbook of Brazilian Conversation*. Chicago, 1943; H. Limas and G. Barroso, *Pequenho dicionário brasileiro de lingua portuguesa*. Rio de Janeiro, 1949.

BretonList (199): F. Vallée, *Vocabulaire français-breton*. St.-Brieuc, 1919.

BretonSE (198): Vannes dialect; M. Piette.

BretonST (198): Standard Breton; M. Piette.

Bulgarian (199): Sofia dialect; Kamen Ganchev.

Byelorussian (199): Eugene Botas.

Catalan (197): José Pujal y Serra, *Diccionario Catalán-Castellano*. Barcelona, 1911.

Creole, see French Creole.

Czech (199): Tatyana Fedorow.

CzechE (199): Trnovec, Slovakia; Alfons Duffek.

Danish (200): Isidore Dyen.

DutchList (199): Isidore Dyen.

EnglishST (200): Modern English.

Faroese (197): Jóhan Hendrik Poulsen, Halgir Winther Poulsen, and Waldemr Dalsgard, via Wayne A. O'Neil.

Flemish (197): L'Abbé Olinger, *Nouveau Dictionnaire de poche français-flamand.* Malines, 1852.

French (200): Isidore Dyen.

FrenchCreoleC (196): Dominican French; Mervyn C. Alleyne.

FrenchCreoleD (196): Dominica, B. W. I.; Douglas Taylor.

Frisian (191): P. Sipma and Y. Poortinga, *Lyts Frysk Wirdboek.* Boalsert, 1944.

GermanST (199): Standard German; Isidore Dyen.

GreekD (198): Demotiki; A. P. D. Mourelatos.

GreekK (198): Katharevousa; A. P. D. Mourelatos.

GreekMD (195): D.C. Swanson, *Vocabulary of Modern Spoken Greek.* Minneapolis, 1959.

GreekML (199): Modern Standard Lesbian Greek; Demetrios Tsenoglou, via E. W. Barber.

GreekMod (198): Modern Greek; Renée Kahane.

Gujarati (194): Naik, via George Cardona.

GypsyGk (183): Greek Gypsy; G. Drachman.

Hindi (200): Frank Southwork.

IcelandicST (195): Standard Icelandic, Rejkjavik and rural dialects; Knut Begsland and Hans Vogt, 'On the validity of glottochronology', *Current Anthropology,* 3.2 (1962): 115-53.

IrishA (199): Miles Dillon.

IrishB (194): T. O. Lane, *Lane's English-Irish Dictionary.* Dublin, 1904.

Italian (199): Isidore Dyen.

Kashmiri (195): G. A. Grierson, *A Dictionary of the Kashmiri Language.* Calcutta, 1932.

Khaskura (187): G. W. P. Money, *Gurkhali Manual.* Bombay, 1942.

Ladin (198): Romanche; Ant. Velleman, *Dicziunari Scurzieu de la Lingua Ladina.* Samaden, 1929.

Lahnda (199): Rishi Gopal Ghatia.

Latvian (199): Riga; K. D. Hramov.

LithuanianO (199): Vitalia Onopiak.

LithuanianST (200): Alfred Senn.

LusatianL (192): Lower Lusatian; Fodor (1961: table 2, opp. p. 334).

LusatianU (192): Upper Lusatian; Fodor (1961: table 2, opp. p. 334).

Macedonian (192): H. G. Hunt, *A Grammar of the Macedonian Literary Language.* Skopje, 1952.

Marathi (200): Frank Southworth.

NepaliList (198): R. L. Turner, *A Comparative and Etymological Dictionary of the Nepali Language*. London, 1931.

[Norwegian] Riskmal, see Riskmal.

Ossetic (184): V. I. Abaev, *Russko-Osetinskij Slovar'*. Moscow, 1950.

PanjabiST (199): Frank Southworth.

PennsylvaniaDutch (191): E. R. Danner, *Pennsylvania Dutch Dictionary*. York, 1951.

PersianList (197): F. Kazemzadeh.

Polish (200): Alexander Schenker.

PortugueseST (200): H. Michaelis, *A New Dictionary of the Portuguese and English Languages*. London, 1893.

Provençal (200): R. P. Xavier de Fourvières, *Lou Pichot tresor, dictionnaire provençal-français-provençal.* Avignon, 1892.

Riksmal (198): Knut Bergsland and Hans Vogt, 'On the validity of glottochronology', *Current Anthropology*, 3.2 (1962): 115-53.

RumanianList (200): Serban Andronescu, *Dictionar de Buzunar Englesz-Romîn.* Bucharest, 1961.

Russian (200): Alexander Schenker.

SardinianC (200): Cagliari dialect; Frederick B. Agard.

SardinianL (200): Logudorese dialect; Dietmar Vogel.

SardinianN (200): Nuorese (Bitti) dialect; Frederick B. Agard.

Serbocroatian (200): Gordana Dimitrijević Lam.

Singhalese (200): B. Anurunddha Vajirarãma, Colombo.

Slovak (200): Tatyana Fedorow.

Slovenian (199): Ljubljana dialect; C. Grote.

Spanish (200): A. Cuyás, *Appleton's New English-Spanish and Spanish-English Dictionary*. New York, 1940.

SwedishList (199): K. Karre *et al.*, *English-Swedish Dictionary*. Stockholm, 1935.

SwedishUp (200): Uppland dialect; Manne Eriksson.

SwedishVL (199): Vilhelmina, Southern Lapland; Rune Vasterlund.

Tadžik (196): D. Arzumznov and X. K. Karimov, *Russko-Tadžikskij Slovar'*. Moscow, 1957.

Takitaki (193): H. R. Wullschlägel, *Deutsch-Negerenglisches Wörterbuch.* Löbau, 1856.

Ukrainian (200): W. Luciw.

Vlach (183): Samarina; G. Drachman.

Wakhi (177): D. L. R. Lorimer, *The Wakhi Language*. London, 1958.

Walloon (198): J. Haust, *Dictionnaire Français-Liègeois*. Liège, 1948.

Waziri (184): John Gordon Lorimer, *Grammar and Vocabulary of Waziri Pashto*. Calcutta, 1902.

WelshC (200): Carmarthen dialect; J. C. Stephens.

WelshN (199): Bangor, North Wales; Hynel Bebb.

References

Anderberg, M. (1973). *Cluster Analysis for Applications*. New York.

Cavalli-Sforza, L. L., Menozzi, P., and Piazza, A. (1994). *The History and Geography of Human Genes*. Princeton.

Cortina Borja, M., and Valiñas Coalla, L. (1989). 'Some remarks on Uto-Aztecan classification'. *International Journal of American Linguistics*, 55: 214-39.

Cox, T. J., and Cox, M. (1994). *Multidimensional Scaling*. London.

Dyen, I., Kruskal, J. B., and Black, P. (1992). *An Indo-European Classification: A Lexicostatistical Experiment*. Philadelphia.

Embleton, S. M.(1986). *Statistics in Historical Linguistics*. Bochum.

Everitt, B. S. and Rabe-Hesketh, S. (1997). *The Analysis of Proximity Data*. London.

Felsenstein, J. (1993). *PHYLIP (Phylogeny Inference Package)*, version 3.5c. Seattle (Department of Genetics, University of Washington).

Fodor, I. (1961) 'The validity of glottochronology on the basis of the Slavonic languages'. *Studia Slavica*, 7: 295-346.

Friedman, J. H., and Rafsky, L. C. (1981). 'Graphics for the multivariate two-sample problem'. *Journal of the American Statistical Association*, 76: 277-87.

Guy, J. B. M. (1980). *Experimental Glottochronology: Basic Methods and Results* (Pacific Linguistics, Series B, 75). Department of Linguistics, Australian National University.

Kaufman, L., and Rousseeuw, P. J. (1990). *Finding Groups in Data*. New York.

Piazza, A., Rendine, S., Minch, E., Menozzi, P., Mountain, J., and Cavalli-Sforza, L. L. (1995). 'Genetics and the origin of European languages'. *Proceedings of the National Academy of Sciences of the USA*, 92: 5836-40.

Kendall, D. G. (1971). 'Seriation from abundance matrices', in F. R. Hodson, D. G. Kendall, and P. Tâutu (eds.), *Mathematics in the Archaeological and Historical Sciences*. Edinburgh, 215-52.

Lohr, M. (2000). 'New approaches to lexicostatistics and glottochronology', in C. Renfrew, A. McMahon, and L. Trask (eds.), *Time Depth in Historical Linguistics*. Cambridge, 209-22.

Renfrew, C. (1992). 'Archaeology, genetics and linguistic diversity'. *Man*, 27: 445-78.

— (1998). *Archaeology and Language: The Puzzle of the Indo-European Origins*. London.

Ross, M. (1997). 'Social networks and kinds of speech community event', in R. M. Blench and M. Spriggs (eds.), *Archaeology and Language, I: Theoretical and Methodological Orientations*. London, 209-61.

Sammon, J. W. (1969). 'A non-linear mapping for data structure analysis'. *IEEE Transactions on Computers*, C-18: 401-9.

Swadesh, M. (1972). *The Origins and Diversification of Language*. London.

Venables, W. N., and Ripley, B. D. (1999). *Modern Applied Statistics with S-PLUS* (3rd edn.). New York.