PUBLICATIONS OF THE PHILOLOGICAL SOCIETY

Verbal Multi-Word Expressions in Corpus Languages

Data, Methodology, and Analysis

Edited by Victoria Beatrix Fendel



VERBAL MULTI-WORD EXPRESSIONS IN CORPUS LANGUAGES

Verbal Multi-Word Expressions in Corpus Languages

Data, Methodology, and Analysis

Edited by Victoria Beatrix Fendel

OpenBook Publishers



https://www.openbookpublishers.com

©2026 Victoria Beatrix Fendel (ed.)
Copyright of individual chapters are maintained by the chapter author(s)





This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). This license allows you to share, copy, distribute and transmit the text; to adapt the text for non-commercial purposes of the text providing attribution is made to the authors (but not in any way that suggests that they endorse you or your use of the work). Attribution should include the following information:

Victoria Beatrix Fendel (ed.), *Verbal Multi-Word Expressions in Corpus Languages: Data, Methodology, and Analysis*. Cambridge, UK: Open Book Publishers, 2026, https://doi.org/10.11647/OBP.0478

Further details about CC BY-NC licenses are available at https://creativecommons.org/licenses/by-nc/4.0/

All external links were active at the time of publication unless otherwise stated and have been archived via the Internet Archive Wayback Machine at https://archive.org/web

Digital material and resources associated with this volume are available at https://doi.org/10.11647/OBP.0478#resources

Information about any revised edition of this work will be provided at https://doi.org/10.11647/OBP.0478

This volume is part of a long-lasting series which can be accessed at https://philsoc.org.uk/Monographs

Publications of the Philological Society vol. 4

ISSN Print: 0265-0649 | ISSN Digital: 2977-845X
ISBN Paperback: 978-1-80511-653-0
ISBN Hardback: 978-1-80511-654-7
ISBN Digital (PDF): 978-1-80511-655-4
ISBN HTML: 978-1-80511-657-8
ISBN Digital ebook (epub): 978-1-80511-656-1
DOI: 10.11647/OBP.0478

Cover photo by Matthew T. Ireland, all rights reserved

Design by Jeevanjot Kaur Nagpal

In der Ruhe liegt die Kraft.

To my parents, Andreas and Brigitte, and
to Maria-Luisa († 2024), who left us far too early.

Table of contents

Contributors ix

List of illustrations xiii

1. Prologue: Challenges and opportunities with multi-word expressions in corpus languages 1
Victoria Beatrix Fendel

I. UNDEREXPLORED DATASETS: CORPUS SELECTION, PREPARATION, AND VISIBILITY 11

2. The Herculaneum papyri as a resource for discovering multi-word expressions in an in-group variety (4th to 1st c. BCE) 13
Michael F. McOsker

3. Technical writing across space and time: The challenges of lemmatisation and part-of-speech tagging in the historiographical and medical corpora of (post-)classical Greek (5th c. BCE to 2nd c. CE) 39
Victoria Beatrix Fendel and Matthew T. Ireland

4. Old Egyptian multi-word expressions in Unas' and Teti's Pyramid Texts (24th to 23rd c. BCE) 73
Roberto A. Díaz Hernández

II. METHODOLOGICAL NOVELTIES: CATEGORIES, CONTINUA, AND THEIR EXPANSION 97

5. Assessing verb-particle constructions through classical Greek preverbation (8th to 4th c. BCE) 99
Elena Squeri

6. The pragmatic development of ποίησον ἀγάπην *poíēson agápēn* as an intersubjective marker in post-classical Greek (4th to 6th c. CE) 131
Lucía Madrigal Acero

7. Between word and phrase: The puzzle of Coptic light-verb constructions (2nd to 5th c. CE) 159
Nina Speransky

III. CRITICAL ENGAGEMENT: MORPHOSYNTACTIC, SEMANTIC, AND PRAGMATIC PERSPECTIVES 187

8. Roundabout ways of setting up scenes in the Greek New Testament (1st to 2nd c. CE) 189
Cressida Ryan

9. Support-verb constructions with causative force in the Vulgate and Septuagint and their Hebrew and Greek originals (3rd c. BCE to 2nd c. CE) 211
Roland Hoffmann

10. Understanding the expression of pain across time: Metaphor and intensity in Latin (and Spanish) support-verb constructions with *dolor* (2nd c. BCE to 15th c. CE) 249
María Isabel Jiménez Martínez

11. Epilogue: A brave new world beyond the canon 283
Victoria Beatrix Fendel

Index 287

Contributors

Victoria Beatrix Fendel (D.Phil. Oxford, 2018) is a researcher in historical linguistics at the University of Oxford, editor of the Classics section of the *Literary Encyclopaedia* and the *Registered Reports in Linguistics*, and language leader for Ancient Greek in the PARSEME (UniDive COST) initiative (www.ancientgreekmwe.com). Her research focusses on language contact and multi-word expressions in literary, epigraphic, and papyrological sources. She is the editor of *Support-verb Constructions in the Corpora of Greek: Between Lexicon and Grammar?* (Language Science Press, 2024) and the author of *Giving Gifts and Doing Favours: Literary Classical Attic Greek Support-verb Constructions* (Brill, 2025).

Roberto A. Díaz Hernández (Ph.D. Tübingen, 2013) is a junior lecturer in Ancient History at the University of Jaén. He is initiator and developer of the Egyptian-UJaen Treebank in Universal Dependencies, the first morphosyntactic treebank of Egyptian, and language leader for Egyptian in the PARSEME (UniDive COST) initiative. His current research focusses on language change in Egyptian and computational linguistics applied to Egyptian. He has recently published *Evidence of Diglossia and Bilingualism in Old Egyptian Mortuary Literature of the Middle Kingdom* (Harvard Egyptological Studies 27, 2025) and *Old Egyptian Multiword Expressions Consisting of a Head Word and ib "Heart"* (Lingua Aegyptia 32, 2024).

Roland Hoffmann (Ph.D. Mainz, 1994) is a researcher in Latin linguistics, editor of the series *Studien zur lateinischen Linguistik* (Buske), and member of the *Mainzer Linguisten-Zirkel*. His research focusses on Latin morphosyntax with a specific focus on Latin support-verb constructions, the history of grammar, and language typology. He is the author of *Lateinische Verbalperiphrasen vom Typ «amans sum» und »amatus fui»: Valenz und Grammatikalisierung* (Peter Lang, 1997) and *Lateinische Linguistik: Morphosyntax und Syntax in einzelsprachlicher und typologischer Perspektive* (Helmut Buske, 2018).

Matthew T. Ireland (MA Cambridge, 2014) is a computer scientist. He is managing director of the company AAI Robotics Ltd and a fellow of Sidney Sussex College, Cambridge. In industry, Matthew works on RF instrumentation

and high-performance compute, specialising in the hardware/software interface and digital signal processing. Matthew also has interests in the digital humanities, having worked on the automated analysis of musical works and automated tagging of linguistic corpora.

María Isabel Jiménez (Ph.D. Universidad Complutense de Madrid, 2016) is Assistant Professor at the Universidad Autónoma de Madrid. Her research explores the semantic, syntactic, and lexical properties of transitive verbs, with a particular focus on verb-noun collocations in Latin and in the early stages of the Romance languages, approached from the functional and typological perspectives. She is the co-editor of *Las colocaciones en la lingüística teórica y aplicada: de las lenguas clásicas a las lenguas romances* (SEEC-Escolar, 2022). Her recent publications include "A Constructional Approach to Causative Support Verbs in Spanish" (co-authored with Chantal Melis), published in *Constructions in Spanish* (John Benjamins, 2023).

Lucía Madrigal Acero (MA Valladolid, 2022) is a predoctoral researcher at the Universidad Complutense de Madrid. Her dissertation focusses on support-verb constructions with ποιέω *poieō* in post-classical Greek. She has collaborated on the *Dictionary of Latin Collocations* (https://dicolat.iatext.ulpgc.es/) and the *Dictionary of Ancient Greek Collocations* (https://dicogra.iatext.ulpgc.es/), coordinated by Baños and Jiménez López. She is a research associate for the project *Interacción del léxico y la sintaxis en griego antiguo y latín 2: Diccionario de Colocaciones Latinas (DiCoLat) y Diccionario de Colocaciones del Griego Antiguo (DiCoGrA)* (PID2021-125076NB-C42) funded by the Spanish Ministry of Science, Innovation and Universities (PRE2022/102525).

Michael F. McOsker (Ph.D. Michigan, 2015) is a researcher in papyrology at University College London in the 'Hexameters Beyond the Canon' project. He was previously an Alexander-von-Humboldt postdoctoral research fellow at the University of Cologne. His scholarship has focussed on the Herculaneum papyri, Epicureanism, and Greek prose stylistics. He is the author of *On the Good Poem According to Philodemus* (Oxford University Press, 2021) and, with David Armstrong (UT-Austin), editor and first English translator of Philodemus' *On Anger* (SBL Press, 2020). He is also a member of the papyrological team for the Vesuvius Challenge.

Cressida Ryan (Ph.D. Nottingham, 2010) has been a Classics teacher and outreach officer, and New Testament Greek lecturer, at the universities of Cambridge, Nottingham, and Oxford. She works on Greek tragedy, the New Testament, and the relationship between these, particularly in the reception of both in the early modern period. Her monograph *Translating the New Testament in Reformation Germany: Reception, Methods, Luther's Latin and Greek Tragedy* will come out

with Amsterdam University Press in 2026. She has further research interests in Neo-Latin, Classics pedagogy, and disability, which also relate to her 2025 roles as disability advisor at the University of Oxford, and fellow at Wolfson College, Oxford.

Nina Speransky (Ph.D. Free University of Berlin, 2023) is a researcher in Coptic and Vietnamese linguistics at the Hebrew University of Jerusalem. Her research focusses on the interplay between morphology, syntax, and semantics. She is the author of *Transitivity and Aspect in Sahidic Coptic: Studies in the Morphosyntax of Native and Greek-origin Verbs* (Widmaier, 2022).

Elena Squeri (Ph.D. Università di Genova/Sorbonne Université, 2023) is a researcher in linguistics at Sapienza University of Rome and a member of the working group for the annotation of Ancient Greek in the PARSEME (UniDive COST) initiative. Her research focusses on the morphosyntax-semantics interface in the Ancient Greek verbal system, on the birth and development of sectorial languages, and on the establishment of grammatical and linguistic concepts in antiquity. She is the author of *La specializzazione semantica nel sistema verbale del greco antico e nella lingua della ginecologia ippocratica* (Schwabe, 2026).

List of illustrations

Figures

10.1	Distribution of the metaphors of *dolor* “pain” in the periods studied. Figure created by the author (2026).	p. 278
11.1	Herculaneum Papyrus number 1251, cornice 1 curated at the Officina dei Papiri in the Biblioteca Nazionale “Vittorio Emanuele III” di Napoli; traditionally identified as Philodemus’ treatise *On Choices* (Περὶ Αἱρέσεων καὶ Φυγῶν *Peri Haireseōn kai Phugōn*). The image is courtesy of Brigham Young University, Provo, Utah and the Ministero per i Beni Culturali. All rights reserved.	p. 284

Tables

3.1	Overview of the number of words, lemmata, and sentences in the sub-corpora.	p. 40
3.2	Error analysis from first-round tagging.	p. 51
3.3	πορεία *poreia* “journey” in Polybius’ *Histories.*	p. 57
3.4	πορεία *poreia* “journey” in the writings by (Thucydides and) Xenophon.	p. 58
3.5	Proximity search for the lemmata χράομαι *khraomai* “to use” and πορεία *poreia* “journey” in the *Thesaurus Linguae Graecae* within a distance of 5 items and manual correction for support-verb constructions.	p. 59
4.1	Distribution of MWEs in Unas’s and Teti’s Pyramid Texts.	p. 75
5.1	Types of change-of-position verbs attested in classical Greek.	p. 113
6.1	Instances of pragmaticalised ποίησον ἀγάπην/ποιήσετε ἀγάπην *poiēson agápēn*/*poiḗsete agápēn* in the *Thesaurus Linguae Graecae* in chronological order.	p. 135

7.1	Sahidic LVCs vs. Bohairic synthetic counterparts in the Biblical corpus.	p. 168
7.2	Bohairic LVCs vs. Sahidic synthetic counterparts in the Biblical corpus.	p. 168
7.3	Sahidic LVCs vs. Bohairic LVCs.	p. 169
8.1	Summary statistics of imperfect verbs and “periphrastic” constructions in the gospels of Mark and Luke.	p. 194
9.1	SVCs of special *verba sentiendi.*	p. 230
9.2	*Dare* “to give” in the three corpora.	p. 231
9.3	Support verbs in the Vulgate.	p. 231
9.4	Position of support verb and abstract noun.	p. 232
9.5	The most frequent SVCs.	p. 233
9.6	All SVCs in the Bible corpus for the verb *facere* “to make”.	p. 236
9.7	All SVCs in the Bible corpus for the verb *dare* “to give”.	p. 237
9.8	Further Hebrew verbs with an SVC in the Vulgate.	p. 240
9.9	The position of the AN and SV.	p. 242
10.1	Distribution of pain SVCs in our corpora.	p. 254
10.2	Distribution of SVCs with *dolor* “pain” in archaic and classical Latin.	p. 257
10.3	Distribution of the metaphors linked with *dolor* “pain” in our corpus of archaic and classical Latin.	p. 264
10.4	Distribution of SVCs with *dolor* “pain” in late Latin and medieval Spanish.	p. 275
10.5	Distribution of the metaphors linked with *dolor* “pain” in our corpus of late Latin and medieval Spanish.	p. 277

1. Prologue: Challenges and opportunities with multi-word expressions in corpus languages

Victoria Beatrix Fendel

Multi-word expressions are multi-lexemic units that function as a single unit from the part-of-speech perspective: for example, English *instead of* functions as a preposition, *event structure* as a noun, *first and foremost* as an adverb, and *to go up* as a verb. This volume is specifically interested in verbal multi-word expressions, such as *to go up*. These take on the role of predicate phrases in the syntax of a sentence, more specifically complex predicate phrases (Bowern 2008: 165).

Structurally speaking, verbal multi-word expressions consist of a verbal component and another component. This other component can be (i) a particle, e.g. in phrasal verbs like *to go up*, (ii) a lexical noun, e.g. in support-verb constructions like *to make a suggestion*, (iii) an adjective, e.g. in causative structures like *to make obvious*, (iv) an adverbial phrase, e.g. in support-verb constructions like *to hold in high esteem*, and (v) a verb, e.g. in serial-verb constructions like *to come see* (see Fendel 2025a: 25). The contributions to this volume (see especially Chapter 3 and Chapter 9) explore what sub-categories of verbal multi-word expressions exist in the languages under investigation as well as how sub-categories that seem to be near-universally available are realised in language-specific ways (Savary et al. 2018: 96; Hoffmann 2023: 29–31; Haspelmath 2010).

Opinions are divided as to whether the components can appear as multiple orthographic words, i.e. separated by whitespaces (for example, Nagy et al. 2013: 329 from a computational perspective), or whether the components can appear as one orthographic word (following univerbation), as in French *maintenir* “to maintain” from Latin *manu tenere* (Lehmann 2020: 217), German *zustandebringen* “to achieve” from **zu Stande bringen* (Peust 1975: 123), or Latin

©2026 Victoria Beatrix Fendel, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.01

animadvertere from *animum advertere* "to pay attention" (Rosén 2020: 266–267). The former perspective is synchronically focussed and raises the issue of structural discontinuity, for example, *to make the question obvious* (Savary et al. 2018: 123; Fendel 2024a), whereas the latter perspective adopts a diachronic view. The contributions to this volume take various perspectives on this issue (see especially Chapter 5 and Chapter 7).

Semantically speaking, verbal multi-word expressions can be compositional or non-compositional, for example, *to take heart* as compared to *to make a suggestion*, and transparent or non-transparent, for example, *to spill the beans* as compared to *to kick the bucket* (Sheinfux et al. 2019: 42 and 50; Mel'čuk 2023: 29–32 and 53; Nunberg et al. 1994). Opinions are divided over whether compositionality is a binary choice (Mel'čuk 2023: 29–32) or operates on a continuum (for example, Langer 2004; 2005). Opinions are also divided over whether transparency is a "psychological property" (Mel'čuk 2023: 53) or part of the semantics (Sheinfux et al. 2019: 42). Verbal multi-word expressions are ambiguous, i.e. between a multi-word-expression and a structurally and semantically compositional reading (Fendel 2025b), and variable, i.e. lexically, morphologically, morphosyntactically, syntactically, semantically, pragmatically, and the list goes on (for example, Savary et al 2018: 123; Sag et al. 2002: 3). The contributions of to this volume explore both the exploitation of such ambiguity for context-specific purposes (see especially Chapter 4) and the diachronic potential such variability proffers (see especially Chapter 6 and Chapter 10).

The language varieties which this volume focusses on are only attested in written form at this point in time. They are not "dead"—for example, Coptic is still used as the language of the Coptic orthodox church (see Ayad 2024)—or "ancient"—for example, Greek and Spanish boast continuous traditions of diachronic development (for example, Horrocks 2014) –, they are simply different. Their native speakers are their corpora of texts (Fleischman 2000: 34). This is why the term "corpus language" has been chosen here. In order to carry out tests for, for example, grammaticality or idiomaticity, we need to ask *them*—in the way that we can. We cannot invite them to attend an interview session at a laboratory, but we can observe them, interact with them, and learn from them (see for example, Vereeck et al. 2023; Wang et al. 2023). We have to adapt to the sampling population we can access, even if it is unbalanced and non-experimental (yet see Tagliamonte 2025: 153). We have to adapt the way in which we operationalise concepts (cf. Whyatt et al. 2025: 30; Ahrens and Janikowski 2025: 189), ask questions (Halliday 1992: 64; Viti 2024: 20–27) and analyse the data (see also Tagliamonte 2025: 110). This is what this volume is about.

It has three main aims, which all the contributions subscribe to and work towards achieving:

1. enhancing the visibility of datasets, both physical and virtual ones, by showcasing the methodologies for doing so (see especially Part I);
2. promoting the interoperability of methodological approaches, through transparent definitions and guidelines (see especially Part II); and
3. furthering the replicability of analyses, through empirically founded approaches with interoperable datasets (see especially Part III).

While the chapters, especially in Part III of the volume, offer interpretative approaches building on the analysis proffered, the volume itself does not focus on the interpretation stage (i.e. the laws of logic, critical thought, hot cognitive skills, etc.) but on the data-focussed pre-interpretation stages. For this reason, each chapter comes with an open-access dataset (and in some cases a codebase) so as to enable the reader not only to replicate the results presented but also to further build on the datasets developed to advance the field.

We cover approaches ranging from computational ones (for example, virtual unrolling of papyri, treebanking of textual corpora, and building Sketch Engine corpora), through linguistic-typological ones (for example, the definition of verb-particle constructions in classical Greek and light-verb constructions in Coptic), sociolinguistic ones (for example, politeness and in-group markers in (post-) classical Greek), philological ones (for example, terminological choices amongst the Epicureans of Herculaneum and translation-language phenomena in the Septuagint and Vulgate), to narratological ones (for example, the eschatological consequences of the framing of scenes in the New Testament). Typologically speaking, Indo-European, i.e. (post-)classical Greek, (post-)classical Latin, and medieval Spanish, and Afro-Asiatic languages, i.e. Egyptian (Old Egyptian and Coptic) and biblical Hebrew, are dealt with. The contributions cover a time period of four millennia starting with the texts by Unas and Teti from Old-Kingdom Egypt (third millennium BCE) and finishing with texts from medieval Spain (fifteenth century CE). Geographically, our texts come from the area around the Mediterranean Sea and adjacent areas.

We look at very different kinds of corpora. The biblical texts (Old and New Testaments, Septuagint and Vulgate) and the Homeric epics (*Iliad* and *Odyssey*) are corpora that exist in a multitude of manuscripts across time (see Ehrman 2006; West 1973, 2015) and have been well digitised and made accessible and searchable in virtual settings (for example, the *Thesaurus Linguae Graecae*, the *Perseus Digital Library*, the *Latin Library*, the *Thesaurus Linguae Latinae*). Conversely, the medical corpora associated with Hippocrates and Galen were long confined to the in-group of medical practitioners (Doody et al. 2012) and the Herculaneum papyri slept a *Rapunzel* sleep under the ashes of Mount Vesuvius and are still relatively confined to an in-group of researchers. These scenarios have brought with them less visibility and accessibility of relevant texts in virtual

settings too. Other corpora are prominent in specialised disciplines, for example, the Pyramid Texts are one of the canonical texts which every Egyptology student will read, and the Byzantine *vitae* will be familiar to historians of Byzantium or theologians. In an effort to open up corpora that have not enjoyed wide-spread interest, visibility, and accessibility, throughout the volume references to textual sources have been given in unabbreviated form, i.e. the discipline-internal abbreviations have been resolved.

The chapters in Part I all revolve around underexplored datasets. Chapter 2 (by McOsker) introduces the corpus of the so-called Herculaneum papyri, named after the place where they were discovered (accessible via the *Thesaurus Linguae Graecae* and papyri.info). The papyri were buried under the ashes of Mount Vesuvius in the first century CE, which carbonised them such that they survived, yet in rolled up form. Technological advancements make it possible nowadays to virtually unroll them without destroying the material. The texts relate to a school of Epicurean philosophers active between the fourth and first centuries BCE. The chapter illustrates the use of multi-word expressions by the in-group in key terminological items. Chapter 3 (by Fendel and Ireland) introduces the corpora of medical writings associated with the physicians Hippocrates and Galen, dating from the fifth century BCE to the second century CE, along with the historiographical corpora by Thucydides and Xenophon (fifth and fourth centuries BCE) and Polybius (second to first centuries BCE) (datasets: https://doi.org/10.5281/zenodo.18440945; https://doi.org/10.5281/zenodo.18441284; https://doi.org/10.5287/ora-jbavbzvz0; codebase: https://github.com/MatthewIreland/xml_lemmatiser_tagger#). While the medical corpora have long been confined to the in-group of medical practitioners, the historiographical corpora became canonical already in antiquity. This has impacted on their visibility, including in digital settings. The chapter introduces the Sketch Engine corpora built for both sets of corpora and illustrates their usefulness in the study of subject- and group-specific multi-word expressions. Chapter 4 (by Díaz Hernández) introduces the treebank built for the so-called Pyramid Texts related to kings Unas and Teti (third millennium BCE) (UniDive release 2.17, November 2025 https://universaldependencies.org/). The Pyramid Texts are utterances that were found on the walls of the burial chambers in the kings' tombs, the pyramids. The chapter illustrates how the Universal Annotation Guidelines for Multi-Word Expressions developed in the PARSEME initiative, originally for spoken languages yet adapted for pre-modern Greek too, can be applied to the Old Egyptian Pyramid Texts. Furthermore, this focus on multi-word expressions in a treebank addresses the frequently observed issue of multi-word expressions being a "pain in the neck for N[atural]L[anguage]P[rocessing]" (Sag et al. 2002; Fendel 2024b). Together, the chapters illustrate three approaches to underexplored datasets in order to make them accessible and visible in virtual

settings as well as searchable and interoperable.

The chapters in Part II all revolve around methodological novelties. Chapter 5 (by Squeri) homes in on the issue of whether multi-lexemic units need to consist of multiple orthographic words, like phrasal verbs, or can form one orthographic word, like preverb-verb combinations in classical Greek (dataset: https://doi.org/10.5281/zenodo.18340160). The history of preverb-verb combinations in Greek is complicated, not least due to the question of tmesis, i.e. the appearance of preverb and verb as two separate units in archaic Greek (eighth to sixth centuries BCE) in particular (for example, Bertrand 2014), and the absence of such tmesis in the Mycenaean tablets dating to the pre-archaic period (ca. fifteenth to twelfth centuries BCE) (for example, McConnell 2025: 130 on the Horrocks vs. Morpurgo-Davies debate). Squeri links this age-old question to the category of verb-particle constructions as defined in the Universal Annotation Guidelines for Multi-Word Expressions developed in the PARSEME initiative, thus querying the language-specific realisation of a comparative concept. Chapter 6 (by Madrigal Acero) provides a detailed study of the semantic and pragmatic polyfunctionality of one specific verbal multi-word expression, ἀγάπην ποιέω *agapēn poieō* "love, favour" and "to do", in fourth- to sixth-century hagiographical texts (dataset: https://doi.org/10.5281/zenodo.17491996). Not only does she illustrate the semantic and pragmatic range of this multi-word expression at any one point in time but also the semantic and pragmatic evolution of this multi-lexemic unit over time, without it univerbating. Chapter 7 (by Speransky) also takes up the issue of whether multi-lexemic units form orthographic words or not, but in a language where the research tradition on how to define wordhood is fundamentally complicated, that is Coptic (see for example, Miyagawa 2024). Her focus is on light-verb constructions, that is, combinations of a verb and a noun, in Coptic literary (including biblical) texts. The verb in such phrases usually appears in the so-called construct state, that is a phonetically reduced form (dataset: https://doi.org/10.5281/zenodo.20330827). The noun can be prototypically eventive or reconceptualised as such and cranberry items appear, for example, ⲁϩⲟⲙ *ahom* in ϥⲓ-ⲁϩⲟⲙ *fi-ahom* (Bohairic dialect) "to utter a sigh". Moreover, morphological innovations (in so-called stative forms) seem to indicate that light-verb constructions form a phonological word in Coptic, for example, ⲕⲁ-ⲣⲱ= *ka-rô*= "to keep silence" with the stative ⲕⲁⲣⲁⲉⲓⲧ *karaeit* formed in analogy to pairs such as ϩⲕⲟ *hko* "to be hungry" and ϩⲕⲁⲉⲓⲧ *hkaeit*. Together, the chapters illustrate the complexity of defining verbal multi-word expressions, the language- and variety-specific realisation of comparative concepts, and the synchronic ambiguity and diachronic developmental potential inherent in verbal multi-word expressions of different types and across typologically unrelated languages.

The chapters in Part III all revolve around critical analysis. Chapter 8 (by

Ryan) takes an eschatological perspective on the use of combinations of a participle with the verb "to be" in the imperfect tense in scene-setting in the Greek New Testament (first and second centuries CE) (dataset: https://doi.org/10.5281/zenodo.18510208). Structures like this have often been dealt with under the label "periphrastic" (for example, Bentein 2016) and the being verb has variously been questioned as a support verb, given its morphosyntax (Jiménez López 2021; Fendel 2026). Ryan touches upon her experimenting with an AI tool (Gemini) in order to draw out relevant instances. However, neither accuracy nor replicability lived up to the bar of this volume (for example, Thomas et al. 2026). Ryan's chapter shows how multi-word expressions are used at narratological turning points in order to guide the reader's attention. They are carefully chosen devices, selected for their semantico-pragmatic properties. Chapter 9 (by Hoffmann) takes a comparative perspective on the Greek and Hebrew texts of the Old Testament and the Greek and Latin texts of the New Testament with a focus on causative support-verb constructions, such as German *Angst machen* "to scare" (dataset: https://doi.org/10.5287/ora-0zvrz5vnk). The chapter devises both a lexeme-based extraction method for relevant instances (i.e. focussing on structures with "to give" and "to make" as well as nouns referring to fearing) and a translation-focussed extraction method (i.e. focussing on German causative constructions in the translation). The chapter is specifically interested in the incidence of multi-word expressions as compared to simplex verbs (for example, the Hebrew Hif'il stem of and the Greek -ίζω *-izō* suffix on verbal roots) and how language-specific differences are dealt with in the translation setting (cf. Riel 2017; Pontani 2023; Gray 2025). Chapter 10 (by Jiménez) takes its starting point from conceptual metaphors. It traces the development of conceptual metaphors that appear with the lexeme *dolor* "pain", when used as a predicative noun in support-verb constructions, through almost two millennia from archaic Latin to medieval Spanish (second century BCE to fifteenth century CE) (dataset: https://doi.org/10.5281/zenodo.19055091). It shows, on the one hand, the temporal depth of multi-word expressions with *dolor* "pain" and, on the other hand, their diachronic variability including both continuities and changes over the two millennia of interest, with a specific focus on the semantico-pragmatic function of intensification. Together, the chapters illustrate the functioning of verbal multi-word expressions beyond the lexico-grammar and specifically in the semantico-pragmatic sphere and thus illustrate their variability and polyfunctionality.

Each chapter is accompanied by an abstract in English and a second, pragmatically preferred language by the author. As will become evident in the chapters of this volume, several corpora, methodologies, and analytic approaches seem to be attached to discourses in one modern language or the other. In the interest of enhancing visibility across these research traditions, we have opted for multilingual abstracts.

At the end of this prologue, it remains to say thank you to all authors, reviewers, and members of the PARSEME GRC working group (Elena Squeri, Rome, Paraskevi Platanou, Athens, Eirene Vlastou, Athens, and Nia Morris, Liverpool). My thanks also go to Agata Savary (Paris), Philomen Probert (Oxford), and James Clackson (Cambridge) for lending a helping hand in a variety of ways and to Roger Macfarlane (Brigham Young University) for making images of the Herculaneum papyri possible in this volume.

Acknowledgements

This work received support from the CA21167 COST action UniDive, funded by COST (European Cooperation in Science and Technology).

Bibliography

Ahrens, Barbara, and Przemysław Janikowski. 2025. Speech Recording. In Ana María Rojo López and Ricardo Muñoz Martín (eds.), *Research Methods in Cognitive Translation and Interpreting Studies*, 183–212. Amsterdam: John Benjamins. https://doi.org/10.1075/rmal.10

Ayad, Mariam (ed.). 2024. *Coptic Culture and Community: Daily Lives, Changing Times*. Cairo: American University in Cairo Press. https://doi.org/10.2307/jj.809342

Bentein, Klaas. 2016. *Verbal Periphrasis in Ancient Greek: Have- and Be-Constructions*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198747093.001.0001

Bertrand, Nicolas. 2014. On Tmesis, Word Order, and Noun Incorporation in Homeric Greek. In Annamaria Bartolotta (ed.), *The Greek Verb: Morphology, Syntax, and Semantics (Proceedings of the 8th International Meeting on Greek Linguistics, Agrigento, October 1–3, 2009)*, 11–30. Louvain-la-Neuve and Walpole, MA: Peeters.

Bowern, Claire. 2008. The Diachrony of Complex Predicates. *Diachronica* 25(2). 161–185. https://doi.org/10.1075/dia.25.2.03bow

Doody, Aude, Sabine Föllinger, and Liba Taub. 2012. Structures and Strategies in Ancient Greek and Roman Technical Writing: An Introduction. *Studies in History and Philosophy of Science Part A* 43(2). 233–236. https://doi.org/10.1016/j.shpsa.2011.12.021

Ehrman, Bart D. 2006. *Studies in the Textual Criticism of the New Testament*. Leiden: Brill. https://doi.org/10.1163/9789047409175

Fendel, Victoria. 2024a. Celebrating Diversity: The Origins and Pathways of Three Support-verb Constructions. *Lexis* 24. https://doi.org/10.4000/12cvp

— 2024b. Multi-word Expressions in the Early Imperial Inscriptions of the I.Sicily Corpus – Still a Pain in the Neck? *Crossreads: Text, Materiality and Multiculturalism at the Crossroads of the Ancient Mediterranean*. 9 October. https://crossreads.web.ox.ac.uk/article/multi-word-expressions

— 2025a. *Giving Gifts and Doing Favours: Literary Classical Attic Greek Support-verb Constructions.* Leiden, Boston, MA: Brill. https://doi.org/10.1163/9789004730526

— 2025b. Support-verb Constructions in the Magical Papyri of the Later Roman and Early Byzantine Periods: A Special Case of Technical Writing. *Symbolae Osloenses* 98(1). 106–136. https://doi.org/10.1080/00397679.2024.2428117

— 2026. From Being, Becoming, and Existing: (Dis)connections between Lexical, Semi-lexical, and Copular Uses of the Being-verb Paradigm. In Ricardo Alcocer Urueta (ed.), *Linguistic Insights into the Meaning and Function of the Verb εἰμί in Classical Greek Philosophy*. Berlin: De Gruyter Brill.

Fleischman, Suzanne. 2000. Methodologies and Ideologies in Historical Linguistics: On Working with Older Languages. In Susan Herring, Pieter Reenen, and Lene Schøsler (eds.), *Textual Parameters in Older Languages*, 33–58. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.195.03fle

Gray, Christa. 2025. The Translator's Patronage: Principles of Responsible Translation in the Jerome-Rufinus Controversy. *Journal of Late Antiquity* 18(1). 65–87. https://doi.org/10.1353/jla.2025.a953360

Halliday, Michael. 1992. Language as System and Language as Instance: The Corpus as a Theoretical Construct. In Jan Svartvik (ed.), *Directions in Corpus Linguistics: Proceedings of the Nobel Symposium 82, Stockholm 4–8 August 1991*, 61–77. Berlin: Mouton De Gruyter.

Haspelmath, Martin. 2010. Comparative Concepts and Descriptive Categories in Crosslinguistic Studies. *Language* 86(3). 663–687. https://doi.org/10.1353/lan.2010.0021

Hoffmann, Roland. 2023. Latin Support-verb Constructions: A View from Language Typology. In José Miguel Baños, María Dolores Jiménez López, María Jiménez Martínez, and Cristina Tur (eds.), *Collocations in Theoretical and Applied Linguistics: From Classical Languages to Romance Languages*, 21–56. Madrid: SEEC.

Horrocks, Geoffrey. 2014. *Greek: A History of the Language and its Speakers*. 2nd edn. Chichester: Wiley Blackwell. https://doi.org/10.1002/9781444318913

Jiménez López, María Dolores. 2021. Γίγνομαι as the Lexical Passive of the Support Verb ποιέω in Ancient Greek. In Georgios Giannakis, Luz Conti, Jesus de la Villa, and Raquel Fornieles (eds.), *Synchrony and Diachrony of Ancient Greek: Language, Linguistics and Philology*, 227–240. Berlin; Boston: Mouton De Gruyter. https://doi.org/10.1515/9783110719192-018

Langer, Stefan. 2004. A Linguistic Test Battery for Support Verb Constructions. *Lingvisticæ Investigationes* 27(2). 171–184. https://doi.org/10.1075/li.27.2.03lan

Langer, Stefan. 2005. A Formal Specification of Support Verb Constructions. In Daniel Schnorbusch and Stefan Langer (eds.), *Semantik im Lexikon*, 179–202. Tübingen: Gunter Narr.

Lehmann, Christian. 2020. Univerbation. *Folia Linguistica Historica* 41(1). 205–252. https://doi.org/10.1515/flih-2020-0007

McConnell, Tom. 2025. *Chronology, Dialect, and Style in Early Greek Hexameter Poetry.* Oxford: Oxford University Press. https://doi.org/10.1093/9780198932345.001.0001

Mel'čuk, Igor. 2023. *General Phraseology: Theory and Practice*. Amsterdam: John Benjamins. https://doi.org/10.1075/lis.36

Miyagawa, So. 2024. Analyticity and Syntheticity in Coptic: Noun Incorporation, Word Eegmentation, and Clitics. In Victoria Fendel (ed.), *Support-verb Constructions in the Corpora of Greek: Between Lexicon and Grammar?*, 291–326. Berlin: Language Science

Press. https://doi.org/10.5281/zenodo.14017939

Nagy, István, Veronika Vincze, and Richárd Farkas. 2013. Full-coverage Identification of English Light Verb Constructions. In Ruslan Mitkov and Jong C. Park (eds.), *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, 329–337. Nagoya: Asian Federation of Natural Language Processing.

Nunberg, Geoffrey, Ivan Sag, and Thomas Wasow. 1994. Idioms. *Language* 70. 491–538.

Peust, Carsten. 1999. Egyptian Phonology: An Introduction to the Phonology of a Dead Language. Göttingen: Peust and Gutschmidt.

Pontani, Filippomaria. 2023. "Faithful" and "Unfaithful" Translations: The Greco-Latin Tradition in Jerome's Letter to Pammachius (395/396 CE). In Glenn Most, Dagmar Schäfe, and Mårten Söderblom Saarela (eds.), *Plurilingualism in Traditional Eurasian Scholarship*, 329–338. Leiden: Brill. https://doi.org/10.1163/9789004527256_029

Riel, Gerd Van. 2017. Augustine's Plato. In Harold Tarrant, François Renaud, Dirk Baltzly, and Danielle Layne (eds.), *Brill's Companion to the Reception of Plato in Antiquity*, 448–469. Leiden: Brill. https://doi.org/10.1163/9789004355385_026

Rosén, Hannah. 2020. Composite Predicates in the Layers of Latin. *Journal of Latin Linguistics* 19(2). 231–279. https://doi.org/10.1515/joll-2020-2009

Sag, Ivan, Timothy Baldwin, Francis Bond, and Ann Copestake. 2002. Multiword Expressions: A Pain in the Neck for NLP. In Alexander Gelbukh (ed.), *Proceedings of the 3rd International conference on Computational Linguistics and Intelligent Text Processing*, 1–15. Berlin; Heidelberg: Springer. https://doi.org/10.1007/3-540-45715-1_1

Savary, Agata, Marie Candito, Verginica Mititelu, Eduard Bejček, Fabienne Cap, Slavomír Čéplö, Silvio Cordeiro, et al. 2018. PARSEME Multilingual Corpus of Verbal Multiword Expressions. In Stella Markantonatou, Carlos Ramisch, Agata Savary, and Veronika Vincze (eds.), *Multiword Expressions at Length and in Depth: Extended Papers from the MWE 2017 Workshop*, 87–147. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.1471591

Sheinfux, Livnat, Tali Greshler, Nurit Melnik, and Shuly Winter. 2019. Verbal Multiword Expressions: Idiomaticity and Flexibility. In Yannick Parmentier and Jakub Waszczuk (eds.), *Representation and Parsing of Multiword Expressions*, 35–68. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.2579035

Tagliamonte, Sali A. 2025. *Analysing Sociolinguistic Variation*. 2nd edn. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009403092

Thomas, Llewellyn D. W., Angelo Kenneth G. Romasanta, and Laia Pujol Priego. 2026. Jagged Competencies: Measuring the Reliability of Generative AI in Academic Research. *Journal of Business Research* 203. 115804. https://doi.org/10.1016/j.jbusres.2025.115804

Vereeck, Alexandra, Mark Janse, Katja De Herdt, Arnaud Szmalec, Cathy Hauspie, and Wouter Duyck. 2023. Why Plato Needs Psychology: Proposal for a Theoretical Framework Underpinning Research on the Cognitive Transfer Effects of Studying Classical Languages. *Psihološka obzorja = Horizons of Psychology.* 121–130. https://doi.org/10.20419/2023.32.580

Viti, Carlotta. 2024. Introduction: Language Contact, Comparative Linguistics, and Comparative Literature in their Historical and Cultural Context. In Carlotta Viti (ed.), *Ancient Greek and Latin in the Linguistic Context of the Ancient Mediterranean*, 11–60. Tübingen: Narr.

Wang, Leonard, Praneet Paidisetty, and Jayme Reisler. 2023. Philoctetes, Pain, and Physicians: Using Greek Tragedy in Medicine. *Ethics, Medicine and Public Health* 28. https://doi.org/10.1016/j.jemep.2023.100887

West, Martin. 1973. *Textual Criticism and Editorial Technique*. Stuttgart: Teubner.

— 2015. History of the Text (HT). In Joachim Latacz and Anton Bierl (eds.), *Prolegomena*, 27–38. Berlin: De Gruyter. https://doi.org/10.1515/9781501501746-005

Whyatt, Bogusława, Anna Hatzidaki, and Brian French. 2025. Participant Profiling. In Ana María Rojo López and Ricardo Muñoz Martín (eds.), *Research Methods in Cognitive Translation and Interpreting Studies*, 21–48. Amsterdam: John Benjamins. https://doi.org/10.1075/rmal.10

I. UNDEREXPLORED DATASETS: CORPUS SELECTION, PREPARATION, AND VISIBILITY

2. The Herculaneum papyri as a resource for discovering multi-word expressions in an in-group variety (4th to 1st c. BCE)

Michael F. McOsker

1. Introduction

In this chapter, I introduce the Herculaneum papyri, a unique collection of primarily Epicurean texts which were carbonised in the 79 CE eruption of Mt. Vesuvius. I explain how they might be useful as a dataset for the study of multi-word expressions in Ancient Greek and describe their editorial situation and most common publication venues. But first, a few words about Epicureanism, the Herculaneum papyri, and the history of scholarship on them.

I am a classicist, more specifically, a papyrologist and historian of ancient philosophy, rather than a linguist; I approach these texts with an eye towards establishing their texts and meanings. They have never been part of the canon, and scholarship on them rarely circulates outside practitioners. Accordingly, this chapter is exploratory and illustrative and more than a little tentative: I hope to make an interesting, coherent group of texts, albeit perhaps a small, strange, and difficult one, more accessible to scholars beyond those who studied them in the past and to give one example of what might be interesting in them, with regard to the study of multi-word expressions.

2. Epicureanism

Epicurus built his philosophy on foundations laid by the shadowy Leucippus and the better-understood Democritus. In short, the Epicureans argued that the only things to exist are atoms, the void, and combinations of these. There are no non-

©2026 Michael F. McOsker, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.02

physical worlds or entities somewhere out there: the Epicureans were rather strict materialists. This had wide-ranging metaphysical and ethical implications: the soul exists, but only temporarily, and it dissolves upon death. Consequently, it cannot be rewarded or tortured in an afterlife. The gods exist, but they live in isolated communities in the *metakosmia*, the spaces between worlds, and do not care about humans one way or the other. Their perfect happiness and peaceful lives are the goal of Epicurean philosophy, but Epicureans engaged in conventional piety without believing that their prayers and sacrifices mattered to gods.

The non-existence of divine rewards and punishments frees the Epicurean to pursue a life of pleasure and the kind of consistent happiness that they called *ataraxia*—literally "freedom from disturbance". The Epicureans were hedonists. Because illegal and immoral actions can produce fear of discovery and punishment, Epicureans are banned from committing them. *Ataraxia* is a long-term goal, not the immediate pleasures of excessive food or drink—which bring future consequences—nor of any illicit gain. Consequently, Epicurus advised his followers to weigh expected pleasures and pains, a process called the "hedonic calculus" by scholars. They must deliberate carefully and choose what to do and not to do. The terms for these decisions are αἱρεῖcθαι *haireîstʰai* "to choose" and φεύγειν *pʰeúgein*, literally "to avoid", but idiomatically "to choose not to do".[1]

3. The Herculaneum library

In 79 CE, Vesuvius erupted and buried Pompeii, Herculaneum, and other towns.[2] Pompeii is the more famous, but Herculaneum turned out to be more important to the history of philosophy. Herculaneum was likely hit by an initial, fast-moving wave of hot ash (hotter than 510 degrees Celsius), before being buried by successive flows of volcanic mud.[3] The hot ash wave carbonised all the organic matter that it touched, such as wooden furniture.

In a gigantic seafront villa just west of town, hundreds of papyrus rolls were found. They had also been carbonised in the ash wave and buried in the eruption, which had the remarkable effect of preserving them for posterity. Organic matter would normally simply decay, as the climate in southern Italy is sufficiently wet. Papyrus normally survives only in very dry climates, such as the Egyptian desert, which is rainless and the source of most ancient papyri.

The villa is now known as the *Villa dei papiri* or "villa of the papyri", from its most famous find, but it was also the source of many important sculptures and

1 For a more detailed introduction to Epicureanism, see Long (1986) and Long and Sedley (1987).
2 On the *Villa dei papiri* and the papyri, the best introduction is Longo Auricchio et al. (2020), of which I am preparing an updated English translation (McOsker in progress).
3 Giordano et al. (2025).

is architecturally interesting in its own right. An unknown number of papyrus rolls were excavated—probably around 600–800. We do not know how many papyri were destroyed before excavators realised what they were, nor how many are still buried, waiting to be discovered.

Initially, the rolls were cut into halves along the long axis, or into quarters, and sketches were made of the visible text. Then, the visible layer would be scraped out or, rarely, lifted out, and the new visible layer would be sketched. This continued until one layer was left, which was preserved. This procedure is called *scorzatura*, or "de-barking", and the resulting pieces are called *scorze* or "pieces of bark". Later, only the outer layers were removed in this way, and the inner part, called the *midollo*, or "marrow", was unrolled more or less continuously on a machine invented by Fr Piaggio, a former librarian in the Vatican. Sketches were also made of these. In both cases, bits or whole stretches of papyrus frequently stick together and are scraped off or unrolled together, disrupting the clean series of columns that would have been legible in the original roll. When these are stuck below the main layer, they are called *sottoposti (*"placed-unders") and when they are above the main layer, they are called *sovrapposti (*"placed-overs"). Sometimes, it seems, entire columns or circumferences of the roll are *sottoposto* to the visible layer. *Sotto-* and *sovrapposti* can often be relocated correctly into their original positions: because they stuck out of place while the roll was in its cylindrical shape, they must always move by multiples of the circumference of the roll, i.e. one or more layers adjacent to their original position.[4]

Piaggio had intended for the *midolli* to be kept as continuous wholes, preserving their original roll-format. However, as this proved impractical, they were cut into smaller sections, attached to heavy paper and ultimately to wooden or stiff cardboard backers. When the Rev. Hayter took all the *disegni* (hand-drawn facsimiles of the papyri) back to England in 1809, a second set was made. The first set is now stored in the Bodleian Library and called the Oxford *disegni*; the second set remained in Naples, in the Biblioteca Nazionale di Napoli, and are called the Neapolitan *disegni*. Because the papyri were not well inventoried before many of them were cut up, the inventory reaches over 1800 entries, but not every entry represents a roll. In many cases, one original roll must be reconstructed out of papyri filed under several inventory numbers. For example, the work tentatively identified as Philodemus' *On Providence* consists of one *midollo* (PHerc. 1670) and four *scorze* (PHerc. 1100, 1577, 1579, and 1636).[5]

The Herculaneum library contains primarily (ca. 90%) Epicurean texts spanning the history of the school from its foundation by Epicurus (341–270)

4 See Nicolardi (2019) for a detailed explanation.

5 See Vergara (2025: esp. 122–144 and the foldout "maquette") for details of the reconstruction of the roll.

down to Philodemus of Gadara (ca. 110–40?). Other texts are Stoic philosophical treatises, Latin literature, perhaps business or legal documents, and perhaps one or two historical works in Greek. Of the Epicurean texts, the majority are by Philodemus, followed by Epicurus.

Thus, the library presents a collection of related texts datable from the late fourth century down to at least the middle of the first century. This unity allows us to follow the development of the school's prose, as well as providing interesting opportunities to study a language community unified by philosophical commitments and by the study of the older texts of their own intellectual tradition, rather than by dialect or time period. But within the corpus of Epicurean texts, we can also study differences by genre, for example, technical treatise, public-facing protreptic, school-internal polemics and letters by famous Epicureans. Thus, the collection can be studied by genre, diachronically, and potentially diatopically: Epicurus spent his career in Athens, along with the other early members of the school, and their texts were presumably composed there. Philodemus, however, spent his career in the company of a Roman senator, and the evidence from the library suggests that he spent time in Southern Italy, in and around the partially hellenophone Naples.[6] Other extant Epicurean texts add to this corpus: some letters by Epicurus (to Menoeceus, Herodotus, and Pythocles) are extant via the usual route of manuscript tradition, and an extensive inscription outlining Epicurean philosophy was erected in Oenoanda by one Diogenes, probably in the second century CE.[7]

4. The difficulties of editing Herculaneum papyri and technological advances

Consequently, editing a Herculaneum papyrus is not easy or straightforward. Firstly, all parts of a roll must be identified. Handwriting and page-layout (i.e. features like the number of lines per column and the length of the lines of text) suggest that pieces may belong to the same roll, but ultimately the plausibility

6 The *Index Academicorum*, a history the Platonic school, exists in Philodemus' draft version (PHerc. 1021 and part of 1691) as well as a finalised version (PHerc. 164); see Fleischer (2023: 79–100) for discussion of the draft version. The Herculaneum library, therefore, descends, at least in part, from Philodemus' own books and working papers.

7 Epicurus' complete letters are available in several editions; the fragmentary ones were most recently edited in Erbì (2020). Diogenes' inscription is best consulted in Smith (1993) and Smith and Hammerstaedt (2014). For further updates, see Smith's series of articles entitled "Diogenes of Oinoanda: News and Notes" that ran in *Cronache Ercolanesi* from a first part in 2006 until part XIV in 2022. See now Smith (2026), a translation of the whole inscription and guide to previous work on it. Diogenes did not use the pair of verbs or of nouns in the surviving part of his inscription, but because he is an Epicurean with extensive surviving texts and because his study is often joined to that of the Herculaneum papyri, it seemed worthwhile to introduce him here.

of the resulting reconstruction confirms it. As a rule, each series of *scorze* must be interleaved with another one (because each series is from half the roll) and they must be read in the opposite sequence of their numbers. Numbers were assigned in the order that *disegni* were created in, but the sketch-artists worked from the visible layer of the *scorze* (the innermost part of the stack) outwards towards the beginning of the roll. The *midollo*, because it was continuously unrolled, can often be read in order, but the order of the pieces of papyrus was open to confusion. The papyri were also hung for a period on museum walls, and many of them had their backing material changed.[8] So, in some cases, the original order of the pieces that constitute the *midollo* must be recovered as well.

The number of sources of text for a particular column or fragment can vary widely. In cases of *scorzatura*, as a rule we have a single *disegno* made before the layer was destroyed. This makes the editor's job easier, because she is not juggling many sources, but the *disegni* have various drawbacks. Their artists often omitted fragmentary or damaged material along breaks and did not usually preserve details about the physical condition of the papyrus: this data can be useful for making joins and determining the sequence of fragments. They were also liable to fail to see or include marginal marks used for punctuation and line numeration, and to confuse similarly shaped letters.

The papyri themselves, when extant, must be consulted. The basic difficulty is that the papyrus itself is darkened, usually now black from carbonisation, so the ink is hardly visible. They are also delicate and liable to flake and crack and cannot be cleaned or conserved like non-carbonised papyri are. They can be read with reasonable success under binocular microscopes with good light, but our ability to read them improved dramatically in the early 2000s, when a Brigham Young University team took infrared photographs of them. These were usually successful in making more text legible. New IR photos, including 3D ones, are being made by the EduceLab team at the University of Kentucky (see https://educelab.engr.uky.edu/).

Therefore, when the papyrus itself is extant, there are usually four main sources for the text: the papyrus itself, two *disegni* (Oxford and Naples), and an infrared photograph.[9] Simple rules cannot be given about the value of the

8 See, in general, Essler (2006). For details, see the individual editions of papyri.

9 In reality, the situation is more complicated. There are often, but not always, earlier scholars' unpublished editions, some dating back to the eighteenth century. The papyri were originally published by engraving the *disegni*, and the proofs and copperplates of these publications occasionally show changes made during the processes of checking the *disegni* before publication. Often, older natural light photographs are extant. These are usually of limited utility but sometimes come in handy if the papyrus was physically damaged after they were taken. Of course, there will also be partial editions and the usual short treatments of specific passages in works on other topics.

sources. The *disegni* were made shortly after unrolling, when the papyrus was in better condition and the ink may have been more visible. However, they are frustratingly vague about the most difficult parts of the text. The infrared photos make the ink more visible but were taken after the papyri had accumulated about 250 years' worth of bangs and bruises. They also flatten the 3D surface of the papyri, which distorts letters standing on curves. The papyri did not unroll flat, but various curves, creases, and wrinkles are burned into them. Most important is the curve that represents the basically cylindrical shape the papyrus was in, but curves may possibly also represent contraction and twisting which took place during carbonisation.

The Vesuvius Challenge's X-ray and AI/machine learning results have made quite a splash in the news recently (see https://scrollprize.org/). The amount of published text from these sources is still small, and their editorial situation is simple: editors work from the images produced by processing the scan data in various ways.[10] *Sotto-* and *sovrapposti* seem to be an issue even in the still rolled-up rolls, either because the pieces are stuck out of place or because the process of digitally unrolling the rolls has not been perfected yet. Less heralded, but also important, is a new computer program, called Maque-IT, which helps scholars rearrange pieces of papyrus, move *sotto-* and *sovrapposti* according to rule, and keep track of the page-layout (see https://recreate-maqueit.it/).

Familiarity makes all this easier, though perhaps never easy. It is important to keep the editorial situation of the papyri in mind, especially the brackets that mark out the letters legible on the papyrus. In (1) below, for example, both verbs that I study in this chapter are partially restored: αἱρο[ύ|μενοι ἢ φεύ]γοντες *hairo[ú]menoi è pʰeú]gontes*. Of these twenty letters, the editors only read ten of them (see below for details); the other ten are restored.

5. "Choosing and avoiding" and "choices and avoidances"

As mentioned at the end of Section 2, the concept of "choice" was important to the Epicureans. Their technical vocabulary for this involves two related pairs of words: αἱρεῖcθαι *haireîstʰai* "to choose" + φεύγειν *pʰeúgein* "to flee", literally, but in this context "to choose to pursue some goal or course of action" and "the choice to avoid or not to pursue one", respectively. The nouns, αἵρεcιc *haíresis* "choice" + φυγή *pʰugḗ* "flight", literally, mean "choice to do something" and "choice not to do something".[11] These meanings have antecedents before the Epicureans (for

10 The sole extensive text so far published can be found in Nicolardi et al. (2024).

11 I have not noticed that any verb is used commonly with the nouns, though they are frequently object of the preposition περί *perí* "about, regarding".

example, in the orator Lycurgus, Plato, and Aristotle, and, contemporaneously, in Stoic authors),[12] but they occur often in Epicurean texts, even as the title of at least two treatises. These technical terms were not unique to the Epicureans, but are very commonly found in their texts, and because of the Herculaneum papyri, we have a lot of Epicurean texts.

I have gathered the passages by surveying the texts (for details of the available resources, see below). What you see is what you get: I have included all passages that I found.[13] Further, I have not limited the list to one part of speech derived from the root (for example, the noun or verb or one of the adjectival formations) but rather included all the pairs of the two roots I came across. In a highly inflected and morphologically flexible language like Ancient Greek, this procedure seems to make sense.[14] It is worth noting that like is found with like: the noun αἵρεσις *haíresis* is found with the noun φυγή *pʰugḗ*, and so forth. Those who wish for a more limited set can prune as they see fit.

I am agnostic about the terminology and definition for these phrases. They are pairs of frequently co-occurring words, whose exact grammatical/formal/lexical realisations vary depending on context, but which are always built to the same pair of roots. These expressions may fall under Laporte's (2018: 176) "non-lexicalised" multi-word expression (depending on how one understands "extra linguistic facts" and his example "salt and pepper"). Fellbaum (2015: 778) uses the term "collocation" for phrases that have a certain "collocational strength", though they are not necessarily fixed; they may be lexically and syntactically idiosyncratic but are semantically decomposable (as opposed to idioms, which are not decomposable). The Epicureans could talk about choices and avoidances separately from each other. PARSEME seems to consider them a nominal idiom.[15]

In my view, the pairing can be reasonably classified as a collocation because the terms—at least within this corpus—cannot be replaced, they cannot be altered or added to, and there's a good argument for understanding it to mean more than the sum of its parts, i.e. "ethical decision making". Since every

12 Lycurgus, *Against Leocrates* 49: τὴν γὰρ δουλείαν φεύγοντες εὐκλεᾶ θάνατον αἱροῦνται *tḕn gàr douleían pʰeúgontes eukleâ thánaton hairoûntai* "for by refusing slavery they chose a glorious death"; Plato, *Republic* 619a, where it may not yet be fully technical; Aristotle, *Eudemian Ethics* 1249b2, Topics 104b7, *Rhetoric* 1360b6, *Nicomachean Ethics* 1173a13. Galen reports uses of the nouns in Chrysippus (*De placitis Hippocratis et Platonis* VIII.2 = SVF vol. III fr. 256.34) and of the verbs in Aristo (*De placitis Hippocratis et Platonis* VII.2 = SVF vol. I fr. 374.5), but these may be Galen's paraphrase.

13 Consequently, I have not registered the dataset or got a DOI for it.

14 Traditionally, the noun αἵρεσις *haíresis*, the verb αἱρεῖσθαι *haireîstʰai*, under the form αἱρέω *hairéō*, the adjective αἱρετός *hairetós* "choice-worthy", and the so-called verbal adjective αἱρετέος *hairetéos* "to be chosen, that must be chosen" are registered separately in the lexica. The same goes for the **pʰeug*- family.

15 See https://parsemefr.lis-lab.fr/parseme-st-guidelines/2.0/index.php?page=030_Categories_of_MWEs/010_Nominal_MWEs

voluntary action is the result of a choice to do or not to do something, the expression "choices and avoidances" encompasses all ethical decision making. This understanding suits the use of the phrase "On Choices and Avoidances" as a title (see below).

In what follows, I give Greek texts and translations only when editions and translations might not be readily available or there is some other reason to quote them. Papyrological texts are laid out according to the papyrus that preserves them, and I use lunate sigma (c) in place of both Byzantine sigmas (σ, ς). Lunate sigma is normal in papyrological texts, but I have extended it to non-papyrological texts for consistency. I have preserved the orthography of the editions, which usually stands in the middle ground between the spellings found in the papyri and modern standards.[16] The papyrological texts are printed with the various signs and marks that papyrologists use to report what they see on the papyrus. The most important of these are square brackets to mark *lacunae* (gaps in the papyrus where nothing can be read), under-dots under letters (ξο̣π̣, to mark uncertainly read letters, i.e. those that can be read as a different letter), and under-dots without letters (to mark either traces of a letter that cannot be deciphered or, when inside square brackets, to indicate space for one letter per under-dot: [. . .]).[17] Double square brackets and high tic-marks (⟦ηθι⟧ and ˋκλμˊ) mark deletions and additions by an ancient scribe. Curly and angle brackets ({αβγ} and ⟨δεζ⟩) mark deletions and additions by the editor.

What follows is a brief collection of passages in Epicurean authors, with the dates and such geographic indications as are available, and works, with brief descriptions of genre and purpose, in which we find either the verbal version or the nominal version of the pair.

The verbal version occurs in Polystratus' *On Irrational Contempt for Common Opinions*, Col. 18 (edited by Indelli), a polemical treatise aimed against sceptical views that denied the reality of observable features, (1).

(1) Polystratus, *On Irrational Contempt for Common Opinions, Col. 18* (ed. by Indelli)[18]

18 . . . .] αἱρο[ύ-

16 For instance, interchange of long iota and the epsilon-iota diphthong is common on papyri and is usually allowed to stand in editions, but otiose or missing *iota mutum* are often "corrected" editorially, e.g. in *On Choices and Avoidances* col. 2.10, (10): οὕτω {ι}. τ[αύτ]ε̣[ι] *hoútō{i} t[aút]e[i]* in Epicurus, *On Nature* 28, fr. 13, col. VIII inf. Sedley, (4), stands for the usual ταύτῃ *taútēi*; this alternative spelling is reasonably common in earlier Epicureans and found very rarely in later ones.

17 A complete introduction to editing papyri can be found here: https://aip.ulb.be/PDF/Guidelines_for_editing_papyri.pdf (note that pp. 30–31 present some of the special sigla used in Herculaneum papyrology).

18 Note that the edition in the *Thesaurus Linguae Graecae* is out of date and that the *Thesaurus Linguae Graecae* column and line numeration does not respect that of the modern edition.

[μενοι ἢ φεύ]γοντες κᾳ[ὶ
20 . . . .] ἐν τῶι βίωι δι-
δᾴ̣ς̣κω[ςι]ν̣ συμφώνω[ς
οἷς] ἡ φύ̣σ̣ις ἡμῶν αὐτ[ὴ
23 [ζη]τεῖ [τ]έ[λ]εσι·
18 *hairo[ú-*
[menoi è pheú]gontes ka[ì
20 *] en tôi bíōi di-*
dáskō[si]n sumfőnōs
hoîs hē phúsis hēmôn aut[è
23 *[zē]teî [t]é[l]esi*

"…choosing or avoiding and … in life let them teach consonantly with the goals our very nature seeks …"

Polystratus was born on Lampsacus but lived most of his adult life in Athens, where he was the third head of the Epicurean school after Epicurus himself and Hermarchus, making him a younger contemporary of Epicurus. Another instance occurs in Philodemus, (2):[19]

(2) Philodemus, *On Rhetoric*, Book 8 = PHerc. 1015, col. 18.5–6 (ed. by Sudhaus, II, p. 11)
2 συμμ[ε]τρο[ῦ]ν[τα] τὸ αἱρε̣[τώ-
τερον λογισμῶι χρῆ[ς]θα[ι
λ. . . πρ. ς, ἐ[πε]ιδή π[ο]τε
5 πρ[οείλη]π̣τ̣α̣ι φεύ[γ]ειν
τι [ἢ αἰ]ρεῖσθα[ι·
2 *summ[e]tro[û]n[ta] tò haire[tő-*
teron logismôi k^{h}rê[s]thai
l. . . pr. s, e[pe]idé̋ p[o]te
pr[oeílē]ptai p^{h}eú[g]ein
5 *ti [è hai]reîst^{h}a[i*

"… the one calculating the more choice-worthy uses reason … since there is a preconception of avoiding something or choosing it."[20]

Philodemus was born in Gadara around 110 BCE and died sometime after 40

19 Note that this text partially repeats text printed at I 294–5 but here it is joined with PHerc. 832; I have used the version printed second. In this text, Sudhaus uses Roman and Arabic numerals to distinguish the two joining papyri.

20 "Preconceptions" (προλήψεις *prolépseis*) are an important part of the Epicurean accounts of knowledge and language use. In short, they are some mental content (a state or image or something like that) that we have which we use to understand the world. Our preconception of a cow lets us identify this thing as a cow and that thing as not a cow (because it is, e.g., a horse).

BCE. He probably left Gadara as a young man and spent a brief period of time in Alexandria before finishing his philosophical education in Athens. In 70 BCE or around then, he joined the entourage of Lucius Calpurnius Piso Caesoninus, a Roman senator and father-in-law of Julius Caesar. He probably resided mostly in Southern Italy, very likely in or around Naples, which was apparently largely Greek-speaking at this date.[21]

The nominal version appears in Epicurus' *Letter to Menoeceus* 128, 129, and 132. The passage at 128 shows clearly the importance of the terms in Epicurean thought, (3).

(3) Epicurus, *Letter to Menoeceus* 128

τούτων γὰρ ἁπλανὴς θεωρία πᾶσαν αἵρεσιν καὶ φυγὴν ἐπανάγειν οἶδεν ἐπὶ τὴν τοῦ σώματος ὑγίειαν καὶ τὴν τῆς ψυχῆς ἀταραξίαν, ἐπεὶ τοῦτο τοῦ μακαρίως ζῆν ἐστι τέλος. τούτου γὰρ χάριν πάντα πράττομεν, ὅπως μήτε ἀλγῶμεν μήτε ταρβῶμεν

toútōn gàr aplanès t^{h}eōría pâsan haíresin kaì p^{h}ugèn epanágein oîden epì tèn toû sṓmatos hugíeian kaì tèn tês psuk^{h}ês ataraksían, epeì toûto toû makaríōs zn̂n esti télos. toútou gàr k^{h}árin pánta práttomen, hópōs mḗte algômen mḗte tarbômen.

"For steady contemplation of these matters knows how to refer every choice and avoidance to the health of the body and the *ataraxia* of the soul, since this is the goal that consists in living blessedly. For we do everything for sake of this, that we feel neither pain nor fear."

Epicurus further uses the terms in *On Nature*, (4):[22]

(4) Epicurus, *On Nature*, Book 28 = PHerc. 1479/1417,[23] fr. [31].[17], l. 20 (ed. by Arrighetti) = fr. 13 col. 8 inf. (ed. by Sedley)

11 [ ]ν κέχρη[ται] τ[αύτ]ε̣[ι] κ[ι]νή-
[cει], καὶ οὐδὲν ἧττον ὑπὸ τὴμ
φυγὴν ἢ αἵρεσιν [κα]τ' αὐτὴν
ἀγόμενος τεύ̣ξεται τοῦ ὁρ-
7 θοῦ.

11 *[ ]n kék^{h}rē[tai] t[aút]e[i k[i]né-*
[sei], kaì oudèn hêtton hupò tèm
p^{h}ugèn è haíresin [ka]t' autèn
agómenos teúksetai toû or-

21 Leiwo (1994).

22 Line numbers in "inf." parts of columns in Sedley's edition count down to 1, the last line in the column. Arrighetti numbers his text within square brackets by text number, then various subsections. Usener (1887) is still a useful collection of fragments related to Epicurus and his doctrine, not fully replaced by Arrighetti (1973).

23 The slash means that the papyri join fairly directly as parts of the *midollo*.

7 *tʰoû*

"… [he] performs this action, and under its guidance, he will arrive at the truth just as much in the category of avoidance as in that of choice."

Epicurus, founder of the Epicurean school, lived from 341–270 BCE. He was born on Samos to an Athenian family and lived briefly in Colophon, Mytilene, and Lampsacus before spending most of his life in Athens. His *Letter to Menoeceus* is a brief protreptic aimed at winning new converts to the school.[24] The *On Nature* was Epicurus' major work, a 37-book treatise covering physics and a number of related topics. At least parts of it (books 28 and what has been identified as 10) were framed as dialogues, and another book (book 25) has dialogue-like elements.[25] However, its overall compositional technique is not yet well understood.

The pair also appears in a fragment of Hermarchus of Mytilene (fr. 34, 8.4 Longo Auricchio), quoted or paraphrased by the much later philosopher Porphyry of Tyre in his *On Abstinence*, a treatise arguing in favour of vegetarianism (I.8.4): ἱκανὴ γὰρ ἡ τοῦ χρησίμου καὶ βλαβεροῦ θεωρία τῶν μὲν φυγὴν παρασκευάσαι, τῶν δὲ αἵρεσιν *hikanὲ gàr hē toû kʰrēsímou kaì blaberoû tʰeōría tôn mèn pʰugὲn paraskeuásai, tôn dè haíresin* "for the contemplation of the useful and harmful is sufficient to produce the avoidance of some things and the choice of others". Hermarchus was a contemporary of Epicurus and became head of the school upon his death in 271 BCE.

The nouns also appear in in several works by Philodemus, (5):[26]

(5) Philodemus, *On Piety*, col. 76

2196 [οἱ δὲ λοι-]
ποὶ διαπ[ίπτουσιν]
ἐν ταῖς αἱρ[έσεσι καὶ]
φυγαῖς, ἱκᾳ[νὰ μηδὲ]
2200 παρ' ἓν ἄγ[οντες]
μήτε κοι[μώμενοι]
τῆι περιφ[ορᾶι· ὁ δὲ]
δίκαιος [περὶ θεῶν ἀ-]
ρίστας ἕ[χει προσδο-]
2205 κίας, ἅμα [δὲ περιτ-]
τῶς καὶ [χρῆται ἀκε-]

24 Menoeceus' identity is unknown, and nothing in the letter is sure to be aimed directly at him.
25 See Lavorante (2024).
26 I reproduce Obbink's (1996) translation.

ραίοις ἡ[δοναῖς κἀπό-]
νοις.
2196 *[hoi de loi-]*
poì diap[íptousin]
en taîs hair[ésesi kaì]
p^{h}ugaîs, hika[nà mēdè]
2200 *par' hèn ág[ontes]*
mḗte koi[mṓmenoi]
têi perip^{h}[orâi; ho dè]
díkaios [perì t^{h}eôn a-]
rístas é[k^{h}ei prosdo-]
2205 *kías, háma [dè perit-]*
tôs kaì [k^{h}rêtai ake-]
raíois hē[donaîs kapó-]
nois.[27]

"… and the rest go astray in their choices and avoidances, neither thinking these things in any respect sufficient, nor resting peacefully on account of disturbance in society with people. But the just person has noble expectations concerning the gods, and at the same time exceedingly enjoys pleasures that are unalloyed and effortless"

Philodemus also mentions the pair, apparently using them as a title, at the end of Περὶ οἰκονομίας, *On Household Management*, (6):[28]

(6) Philodemus, Περὶ οἰκονομίας, *On Household Management*, col. 28.9
28.48 ἀλ-
λ' ὀλο̣χερεστέρων ἀνάγ-
29.1 κην εἶναι ποιεῖ[cθ]α[ι τ]ὰ[c π]αρ[α-]
δόcειc καὶ πολλὰ κατὰ [μέροc ὑ-]
πογράφ[ε]ιν διεξόδο[ιc πρὸc ἐπι-]
μέλειαν καὶ φυλακὴν [χρη]μά-
5 των, καὶ [μ]άλι[c]τ̣' ἐν τῆ[ι περ]ὶ
πλούτου καὶ πενίαc κα[ὶ τῆι]
πε[ρ]ὶ δι[αί]τηc πολυτε[λοῦc]

27 The final word is κἀπόνοις *kapónois*, a crasis of καὶ ἀπόνοις *kaì apónois*.
28 This text was most recently edited by Tsouna (2012), but she reproduces the out-of-date papyrological conventions of the previous edition, Jensen (1907). I have silently updated them. The square brackets in the translation represent damaged Greek words. Also, she accidentally omitted ὑ]|πορά̣φ[ε]ιν *hu]|pograp^{h}[e]in*, 29.2–3 in her Greek text, though she translates it; I have put it back in its right place.

τε καὶ λ[ι]τῆς καὶ τῆι περ[ὶ]
αἱρέσεων καὶ φυγῶν κἄν εἴ
τινές εἰσιν ἄλλαι τοια[ῦ]ται.

all-
l' holosk^{h}erestérōn anág-
kēn eînai poieî[st^{h}]a[i t]à[s p]ar[a-]
dóseis kaì pollà katà [méros hu-]
pográp^{h}[e]in dieksódo[is pròs epi-]
méleian kaì p^{h}ulakèn [k^{h}rē]má-
tōn, kaì [m]áli[s]t' en tê[i per]ì
ploútou kaì penías ka[ì têi]
pe[r]ì di[aí]tēs polute[loûs]
te kaì l[i]tês kaì têi per[ì]
hairéseōn kaì p^{h}ugôn kàn[29] *eí*
tinés eisin állai toia[û]tai

"… it is, nevertheless, necessary [to hand down a tradition] of the most general principles and to outline [many details] in the treatises [concerning] the care and preservation of possessions, and [chiefly in] the treatise on wealth and poverty, the one on the luxurious and the frugal ways of life, the one on things to be chose and to be avoided, and in any other treatises of this sort"

Philodemus' *On Rhetoric* is a fertile source of the pair, (7).[30] This work is interesting because the first three books are devoted to a fight within the Epicurean school over the interpretation of the founding generations' doctrine(s) about rhetoric. Only with books 4–8 (or more) does Philodemus' gaze turn outwards:[31]

(7) Philodemus, *On Rhetoric*, Book 2 = PHerc. 1672, col. 22 (ed. by Longo Auricchio, pp. 216–218)

ἀρκέσει τὰ πα-
ρα Μητροδώρωι σαφῶς διδά-
σκοντα· …
καὶ μικρὸν προβάς· "τί γὰρ [ὑ-]
πόκειται; ὃ θεωρεία π̣[ο]ιεῖ,
ὥσπερ ἡ περὶ τὰς αἱρέσεις

29 A crasis of καὶ ἐάν *kaì eán* "and if".

30 The editorial situation of Philodemus' *On Rhetoric* is ridiculously complicated. It was in at least eight books, parts of at least seven of which survive, but several (books 2–4) do so in two copies.

31 See Longo Auricchio (1985: 36–37) and Chandler (2006: 50), to whom my translation is indebted.

καὶ φυγὰς καὶ περὶ τὰς πολει-
τικὰς ἐκ τῆς ἱστορίας πα-
ρακολουθήσεις;”

aréskei tà pa-
ra Mētrodṓrōi sap^{h}*ôs didá-*
skonta: …
kaì mikròn probás: “tí gàr [hu-]
pókeitai? hò t^{h}*eōreía p[o]ieî*
hṓsper hē perì tàs hairéseis
kaì p^{h}*ugàs kaì perì tàs polei-*
tikàs ek tês historías pa-
rakolout^{h}*éseis?*

“… the bit in Metrodorus that teaches this clearly will suffice…and a little further on: ‘What is the foundation? What investigation produces, just like that regarding choices and avoidances and regarding the political consequences discovered by research?’”

The usage here, it is worth noting explicitly, is quoted by Philodemus from Metrodorus of Lampsacus, Epicurus’ second-in-command who predeceased him, (8) and (9).[32]

(8) Philodemus, *On Anger*, Book 6 (?) = PHerc. 1669, col. 13.26 (ed. by Sudhaus, p. I 245)

[ἐ]ν τ[οῖ]ς τ[ῶν] π[ε-
ρὶ ἠθῶν [κ]αὶ βίων κἀ{ι}π[ὶ
τὰς αἱρέσεις καὶ φυγὰς
διατεινόντων ἀναςτρέ-
φεται, δι’ ὧν καὶ τὰ κατὰ
τὴν πόλιν καὶ τὴν ἀγο-
ρὰν ὀρθότατα πάν[τ]ων
— ἔστιν οἰκονομεῖν.

[e]n t[oî]s t[ôn] p[e-
rì ēthôn [k]aì bíōn kap[ì[33]
tàs hairéseis kaì p^{h}*ugàs*
diateinóntōn anastré-
p^{h}*etai, di’ hôn kaì tà katà*
tèn pólin kaì tèn ago-

32 For the most recent edition of this column, see Di Matteo (2005: 121).
33 κἀπὶ *kapí* is a crasis for καὶ ἐπί *kaì epí*.

30 *ràn orthótata pánt[ō]n*
éstin oikonomeîn.

"… he [sc. the philosopher] is engaged in the questions regarding characters and lifestyles that relate to choices and avoidances, through which it is possible to manage matters related to the city and commerce most correctly of all."

(9) Philodemus, *On Rhetoric*, Book 8 = PHerc. 1015, col. 10.2 (ed. by Sudhaus, I, p. 288)

1 . . . . . ὅτι χρὴ [θε]ωρήσει [διακρῖ-
ναι] τὰ[c] αἱρέσεις καὶ φ[υγὰς
τὰς π]ρὸς [μα]καριό̣τητ[α
συν]τ̣ειν̣ούσας καὶ οὐ τὰς
5 πρὸς [ἐ]λ̣π̣ισμὸν ἀόρι-
στον καὶ π]ρὸς δοξοκ[ο-
π]ίαν ἄλο̣[γ]ον …

1 *. . . . . hóti k^{h}rè [t^{h}e]ōrḗsei [diakrî-*
nai] tà[s] hairéseis kaì p^{h}[ugàs
tàs p]ròs [ma]kariótēt[a
sun]teinoúsas kaì ou tàs
5 *pròs [e]lpismòn aóri-*
ston kaì p]ròs doksoko-
p]ían álo[g]on …

"… that it will be necessary to evaluate by contemplation the choices and avoidances that pertain to blessedness and not those that pertain to unlimited hope and to irrational glory-seeking …"

Lastly, they are paired six times in Philodemus' *On Choices and Avoidances* (Περὶ αἱρεσέων καὶ φυγῶν *perì haireséōn kaì p^{h}ugõn*), (10) to (14). The names of the treatise and its author do not survive on the papyrus, but those seem most likely.[34] The title had been previously used by Epicurus (reported by Diogenes Laertius 10.27), and Philodemus mentions it as a common topic at the end of *On Household Management* (see above).[35] Its use as a title is particularly interesting,

34 See the edition by Indelli and Tsouna-McKirahan (1995). The brackets around the author and ancient title in their modern title indicate that those are scholarly restorations rather than readings of the papyrus. The text used to be referred to as the *Ethica Comparetti*, after an editor. Translations of this text are those of Indelli and Tsouna-McKirahan (1995).

35 Note that in the *Thesaurus Linguae Graecae*, the line numerations of this text follow those of the printed edition, not the papyrological lines, though the upright lines (|) to mark line division in the papyrus are preserved. Thus, the *Thesaurus Linguae Graecae* records the first citation as being in col. 2, line 6, because it is in the sixth line of text in the printed edition, but I call it lines 10–11, because it starts in the tenth papyrological line. Similarly, the *Thesaurus Linguae Graecae* preserves the printed line-end hyphenation of the book, rather than hyphenating at

since titles often draw on established frameworks (note, for example, how common the format “Pop Culture Reference/Joke: Informative Subtitle” is in academic publishing).

(10) Philodemus, *On Choices and Avoidances*, col. 2

5 κα[ὶ κατ]αξι[οῦ-]
cιν ὄν]τωc μηδὲ[ν] προκε[ῖc-]
θαι πεπ]ειcμένοι ὅτ[ι τ]ὸ μέ[γα]
πάθοc τὸ] ψυχικὸν ὑπὲρ τοῦ
[πόνο]υ γίνεται καὶ cυμ-
10 [περαίνο]μεν οὕτω {ι} τὰc αἱ[ρ]έ-
[ceιc κ]αὶ φυγὰc πρὸc ἅμ[φω]
βλέπ]οντεc. οὐδὲ γὰρ [ἦ]ν
[τὰc] χαρὰc ὁμοίωc ἀ[θ]ρό[ac
ἐγγί]νεcθαι κατ’ ἐλπί[δα . . .
15 τιν]ά …

5 *ka[ì kat]aksi[oû-]*
sin ón]tōs mēdè[n] proke[îs-]
thai pep]eisménoi hót[i t]ò mé[ga]
páthos tò] psuk^{h}ikòn hupèr toû
póno]u gínetai kaì sum-
10 *peraíno]men hoútō tàs hai[r]é-*
seis k]aì p^{h}ugàs pròs ám[p^{h}ō]
blép]ontes. oudè gàr [ê]n
tàs] k^{h}aràs homoíōs a[t^{h}]ró[as
eggí]nest^{h}ai kat’ elpí[da…
15 *tin[á…*

“And they claim that in truth no (judgment) takes precedence over any other, being persuaded that the great affection of the soul occurs as a result of pain and that thus we accomplish our choices and avoidances by observing both [sc. bodily plus mental pain]. For it is not possible that the joys arise in us in the same way and all together, in accordance with some expectation…”

line breaks in the papyrus, as is the usual procedure.

(11) Philodemus, *On Choices and Avoidances,* col. 5

μετὰ ḍ[ὲ τ]αῦτ[α καὶ] τὰc ṭῶν
ἐπιθυμ[ιῶ]ν π[ερί] ṭε τὰ[c] ἡδο-
νὰ[c] καὶ ṭὰ ποι[ητ]ικὰ διᾳφᾳ-
ρὰc ἀναλογιc[τέ]ον· ἐπ[ειδὴ]
καὶ παρὰ τὴν [ἀδι]αληψ[ία]ν [δι’
αὐτῶν μεγάλα γίνεται δ[ι-
απτώματα κατὰ τὰc αἱρέc[εις]
καὶ φυγάc. ἕνεκα γὰρ τῶ[ν] ξε-
νοτάτων ὡc ἀναγκαιοτά-
των τὰ χαλεπώτατ’ ἀναδέ-
χ[ο]νται κακά, δυναcτείαc
λέγω καὶ λαμπρᾶc δόξηc
καὶ π[ε]ριουcίαc ὑπεραγού-
cηc καὶ ṭ[ρυ]φῶν τοιούτων
καὶ τῶν ὁμοίω[ν]

metà d[è t]aût[a kaì] tàs tôn
epit^{h}um[iô]n p[erí] te tà[s] hēdo-
nà[s] kaì tà poi[ēt]ikà diap^{h}o-
ràs analogis[té]on: ep[eidè]
kaì parà tèn [adi]alēps[ía]n [di’
autôn megála gínetai d[i-
aptṓmata katà tàs hairés[eis]
kaì p^{h}ugás. héneka gàr tô[n] kse-
notátōn hōs anagkaiotá-
tōn tà k^{h}alepṓtat’ anadé-
k^{h}[o]ntai kaká, dunasteías
légō kaì lamprâs dóksēs
kaì p[e]riousías huperagoú-
sēs kaì t[ru]p^{h}ôn toioútōn
kaì tôn homoíōn

“Having looked into those matters, one should also consider the differences among desires, both with regard to the pleasures and with regard to their causes. For it is also on account of the failure to distinguish between them that important errors occur through them with respect to choices and avoidances. For men suffer the worst evils for the sake of the most alien desires which they take to be the most necessary—I mean desires for sovereignty and brilliant reputation and great wealth and suchlike luxuries and other similar things.”

(12) Philodemus, *On Choices and Avoidances*, col. 11

5 ὃ
[διὰ] τὰ π[ε]ρὶ τῶ[ν] τεττάρω[ν εἰ-]
[ρ]ημέν̣α λέγεται, το[ῦ] τὴ[ν περί-]
λ̣ηψιν τὴν περὶ τῶν κυρι[ωτ]ά̣-
[τ]ων καὶ τὴν μνήμην π̣[ολ-]
10 λὰ συμβάλλεσθαι πρὸς τὰς
οὔσας αἱρέσεις καὶ φυγὰς οὐ-
κ̣ ἴσου τιθεμένου, καθάπερ
ἐξεδέξαντό τινες ἀγροί-
κως, τῶι τινας ἀναφέρ̣ες-
15 θαι τῶν αἱρέσεων καὶ φ̣υ̣γῶν
ἐπὶ τὰς περὶ τούτων ἀταραξίας,
ἀλλὰ τῶι κ[α]τ̣ορθοῦς-
θαι μὲν αὐτὰς τοῖς τέλεσι
τοῖς τῆς φύσεως παραμε-
20 τ[ρ]ούντων …

5 *hò*
[dià] tà p[e]rì tô[n] tettárō[n ei-]
[r]ēména légetai, to[û] tḕ[n perí-]
lēpsin tḕn perì tôn kuri[ōt]á-
[t]ōn kaì tḕn mnḗmēn p[ol-]
10 *là sumbállest*^{h}*ai pròs tàs*
oúsas hairéseis kaì p^{h}*ugàs ou-*
k ísou tit^{h}*eménou, kat*^{h}*áper*
eksedéksantó tines agroí-
kōs, tôi tinas anap^{h}*éres-*
15 *t*^{h}*ai tôn hairéseōn kaì p*^{h}*ugôn*
epì tàs perì toútōn ataraksías,
allà tôi k[a]tort^{h}*oûs-*
t^{h}*ai mèn autàs toîs télesi*
toîs tês p^{h}*úseōs parame-*
20 *t[r]oúntōn*

“This is said because of what has been stated about the four cardinal principles, for the thesis that the understanding and the memorisation of the cardinal tenets contribute greatly to actual choices and avoidances is not equal to claiming that some choices and avoidances are traced back to the states of tranquillity concerning them [sc. the cardinal tenets]—as some have clumsily interpreted

it—but to claiming that they [sc. the choices and avoidances] are accomplished successfully when we measure them by the ends laid down by nature …"

(13) Philodemus, *On Choices and Avoidances*, col. 13

[κ]αὶ δηλονότι καὶ τὰ συγγενικὰ
τέλη προσφερόμενα τὴν ἐ-
γδηλοτάτην ἐνάρ[γ]ειαν, οἷς αἱ
παραμετρή[c]εις γίνονται τῶν
αἱρετῶν καὶ φευκτῶν. καὶ
τοὺς ἠθικοὺς καὶ περὶ τῶν
αἱρέσεων καὶ τῶν φυγῶν λό-
γους προσλαβεῖν δεῖ πάντως
ἐκ φυσιολογίας ἵν' ἐντελεῖς
ὦσι[ν]…

[k]aì dēlonóti kaì tà suggenikà
télē prosp^{h}erómena tèn e-
gdēlotátēn enár[g]eian, hoîs hai
parametré[s]eis gínontai tôn
hairetôn kaì p^{h}euktôn. kaì
toùs ēt^{h}ikoùs kaì perì tôn
hairéseōn kaì tôn p^{h}ugôn ló-
gous proslabeîn deî pántōs
ek p^{h}usiologías hín' enteleîs
ôsi[n] …

"… and it is clear that he also establishes the congenital ends, which yield the most conspicuous evidence and by which the calculations concerning choices and avoidances are performed. Besides, one must unfailingly draw the moral arguments regarding both choices and avoidances from the study of nature in order that they should be complete …"

(14) Philodemus, *On Choices and Avoidances*, col. 14

–1 [οὐκ ἔστιν]
1 [ἡδέωc ζῆν] ἄνευ [τοῦ φρο-]
[νί]μωc κ[αὶ καλ]ῶc [καὶ δικαί-]
ωc, ἔτι δ' ἀνδ[ρ]είωc [κ]α[ὶ ἐγ-]
κρατῶc καὶ μεγαλο[ψύ]χ[ω]c
5 καὶ φιλοπ̣ο̣ητικῶc κα[ὶ φιλαν-]
[θρ]ώπωc καὶ καθόλ[ου πασῶν]
τῶν ἄλλων ἀρετῶν ὑπα[ρ-]

χουcῶν, ὡc τῶν μεγί[cτω]ν
παραπτωμάτων ἐ[ν] ταῖc
αἱρέ⟨cε⟩cι καὶ ταῖc φυγαῖc cυντε-
λουμένων, ὅταν τὰ[ν]αντία
νομίζοντεc ἔνιοι καὶ διὰ το[ῦ-]
το cυνεχόμενοι ταῖc κακί-
αιc ἕκαcτ' ἐνεργῶcιν·

[ouk eśtin]
[hēdéōs zên] áneu [toû p^{h}ro-]
[ní]mōs k[aì kal]ôs [kaì dikaí-]
ōs, éti d' and[r]eíōs [k]a[ì eg-]
kratôs kaì megal[psú]k^{h}[ō]s
kaì p^{h}ilopoētikôs ka[ì p^{h}ilan-]
[t^{h}r]ópōs kaì kat^{h}ól[ou pasôn]
tôn állōn aretôn hupa[r-]
k^{h}ousôn, hōs tôn megí[stō]n
paraptōmátōn e[n] taîs
hairésesi kaì taîs p^{h}ugaîs sunte-
louménōn, hótan ta[n]antía[36]
nomízontes énioi kaì dià to[û-]
to sunek^{h}ómenoi taîs kakí-
ais hékast' energôsin

"(… it is impossible for one to live pleasurably) without living prudently and honourable and justly, and also without living courageously and temperately and magnanimously, and without making friends and without being philanthropic, and in general without having all the other virtues. For the greatest errors in choices and avoidances occur when some men perform each action while holding the opposite views, and thereby while being constrained by evils"

Note also αἱρετὰ καὶ φευκτά *hairetà kaì p^{h}euktá* "choice-worthy and avoidance-worthy things" in Philodemus' *On Rhetoric*, Book 1, col. 40.31 (ed. by Nicolardi), above in Philodemus' *On Choices and Avoidances* col. 13, (13), and several other times in the Epicurean corpus (Philodemus, *History of the Stoics*, col. 31 and *On Poems*, Book 1.80. The pair αἱρετικός *hairetikós* + φευκτικός *p^{h}euktikós* occurs in *On Music* IV col. 138.12–22 Delattre.

This survey includes four authors from the fourth to third centuries BCE (Epicurus, Metrodorus as quoted by Philodemus, Hermarchus, and Polystratus) and from the first century BCE (Philodemus); Philodemus preserves several

36 Another crasis, for τὰ ἐναντία *tà enantía*; see above.

instances of the pairs. All are educated and presumably native speakers of Greek from a few different homelands.[37] They all lived in Athens for extended periods. The earliest ones were relatively early writers of Koinē; Philodemus wrote as Atticism began. Epicureans as a matter of philosophical principle privileged clarity over style in writing, but Epicurus and apparently Metrodorus could write attractively when they wanted to. Philodemus always obeyed at least one important stylistic rule (avoidance of hiatus) in everything he wrote and varied his style in different treatises by occasionally writing rhythmically.[38] There are many interesting variables.

7. Bibliographic guide and advice

The Herculaneum library's publication situation presents some difficulties for researchers. Texts are published overwhelmingly in the traditional print formats (book-length complete editions and journal article-length partial editions), and they are not reliably collected into conveniently digitised corpora. Many can be found in two databases, namely papyri.info and the *Thesaurus Linguae Graecae*. papyri.info is a hub for everything papyrus-related; it incorporates a number of digital projects and includes texts of many papyri, including many of those from Herculaneum. The *Thesaurus Linguae Graecae* is a searchable database of Greek texts from the earliest alphabetic Greek texts (Linear B is not included) until the nineteenth century.

The online inventory *Chartes* is especially useful. It gives title and author identifications, language, conservation information, and notes about matches with other pieces, as well as a handwriting sample and bibliography. It also links to the *Leuven Database of Ancient Books* (LDAB) entry for papyri, which makes it easy to find the papyri.info text.

The situation is made more complicated by the fact that there are many editions of *parts* of texts published in various journals, some of which are specialised and do not circulate widely. This makes it difficult to study a whole text. Take as an example PHerc. 831, thought to be a work by Demetrius the Spartan on false

37 Gadara was heavily Hellenized, and Philodemus gives no indication that he used a second language, besides using a few Latin words in otherwise Greek texts. Meleager, who says he grew up in Tyre and Gadara and who lived a generation or so before Philodemus, claimed to speak Greek, 'Syrian' (a local Aramaic dialect?), and Phoenician, but it is not clear if that was normal for Gadara; cf. *Palatine Anthology* 7.419.7–8: ἀλλ᾽ εἰ μὲν Σύρος ἐσσί, Σάλαμ· εἰ δ᾽ οὖν σύ γε Φοῖνιξ, | Ναίδιος· εἰ δ᾽ Ἕλλην, Χαῖρε· τὸ δ᾽ αὐτὸ φράσον *all᾽ ei mèn Súros essí, Sálam; ei d᾽ oûn sú ge Pʰoîniks, | Naídios; ei d᾽ Héllēn, Kʰaîre; tò d᾽ autò pʰráson* "But if you are Syrian, *salām*, and if you are Phoenician, *naidios*, and if Greek, *khaire*—now you say the same!"

38 See McOsker (2017) on avoidance of hiatus in these authors and McOsker (2022) about Philodemus' rhythmic prose. In a future article, I intend to discuss why he varies his style. Indelli and Longo Auricchio have dedicated a series of six articles—one in each issue of *Cronache ercolanesi* from 50 to 55—to Philodemus' use of *hapax legomena*.

beliefs. The most complete edition is in Koerte's 1890 edition of Metrodorus (he was probably wrong about its authorship), but of the twenty-one columns and fragments he published, seventeen must be read in newer editions published by three different scholars.[39] A would-be user of this treatise must assemble the partial texts from all these publications into a single text file, and the result will be a sort of Frankenstein's monster of editorial choices. papyri.info has a text, but the *Thesaurus Linguae Graecae* does not. The bibliography on Chartes and the "Reference Edition" and "Partial Editions" data on papyri.info make bibliographic hunting easier.

The main venue for publication of Herculaneum texts is the Bibliopolis series *La scuola di Epicuro* ("The School of Epicurus") which had a sub-series dedicated to the fragments of the early leaders of the Epicurean school (*I frammenti dei Kathegemones* "The Fragments of the Leaders") and a supplement series (with editions of the Latin poet Lucretius and the Greek inscription of Diogenes of Oenoanda). The series has twenty-one volumes, with a twenty-second due out soon.[40] The *Society for Biblical Literature's Writings from the Greco-Roman World* series has four editions; the *Les Belles Lettres' Budé* series has two (with more expected), and Oxford University Press has published four fifths of Philodemus' *On Poems* and half of his *On Piety*. A new series with Brill, *Papyri Graecae Herculanenses* has two titles. The most important journal is *Cronache ercolanesi,* the long-running home of Herculaneum studies. Related material can be found in most papyrological journals (especially *Zeitschrift für Papyrologie und Epigraphik*) and the new journal *Philosophical Papyri*. Other journals dealing with papyrological and philosophical topics occasionally publish Herculaneum material.

The role of technology in editing the papyri means that, as a rule, newer editions are overall better than previous ones. More recent editors simply have better equipment than older ones. However, this does not mean that their judgment in supplementing damaged traces and choosing conjectures is better, and older editors often saw the papyri when they were in better condition than they are now. Ideally, a critical reader checks older editions for relevant material while keeping an eye on the sources (papyrus, infrared photos, *disegni*) for the text. One must read the introduction of each text for details of reconstructions. Every papyrus is different and presents unique problems. Likewise, though some uniformity has been enforced by the Leiden conventions, editors still use different conventions for material preserved in various sources and readings

39 Beyond Koerte (1890), see the entries for Schmid, Sanders, and Parisi in the bibliography. Beyond attempts to edit the text, various scholars published reviews of Koerte's book or passing observations in works dedicated to other topics.

40 At least two editions, those of Philodemus' *On Anger* and *On Music* IV, have been superseded by later editions.

changed by the editor, so the sigla at the head of the edition should be checked.

8. Concluding remarks

Despite the difficulties, the number of surviving texts from a single library (and the fact that they mostly belong to the Epicurean school) makes the Herculaneum library an important resource for ancient philosophers. It could be important for other scholars as well. In this chapter, I hope to have demonstrated the interest—and to have alleviated some of the difficulties—that these papyri present.

Bibliography

Arrighetti, Graziano. 1973. *Epicuro: Opere*. 2nd edn. Turin: Einaudi.

Chandler, Clive. 2006. *Philodemus on Rhetoric: Books 1 and 2: Translation and Exegetical Essays*. New York: Routledge. https://doi.org/10.4324/9780203959626

Delattre, Daniel. 2007. *Philodème de Gadara: Sur la musique. Livre IV*, 2 vols. Paris: Les Belles Lettres.

Di Matteo, Tiziana. 2005. Segni di interpunzione nel *PHerc*. 1669: tipologia grafica e funzione. *Cronache Ercolanesi* 35. 119–124.

Essler, Holger. 2006. Bilder von Papyri und Papyri als Bilder. *Cronache Ercolanesi* 36. 103–143.

Fellbaum, Christine. 2015. Syntax and Grammar of Idioms and Collocations. In Tibor Kiss and Artemis Alexiadou (eds.), *Syntax—Theory and Analysis: An International Handbook*, 777–802. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110377408.777

Fleischer, Kilian. 2023. *Philodem: Geschichte der Akademie: Einführung, Ausgabe, Kommentar.* Leiden: Brill. https://doi.org/10.1163/9789004546547_13

Giordano, Guido, Alessandra Pensa, Alessandro Vona, Danilo Di Genova, Raschid Al-Mukadam, Claudia Romano, Joachim Deubemer, Alessandro Frontoni, and PierPaolo Petrone. 2025. Unique Formation of Organic Glass from a Human Brain in the Vesuvius Eruption of 79 CE. *Science Reports* 15, 5955. https://doi.org/10.1038/s41598-025-88894-5

Hammerstaedt, Jürgen, and Martin Fergusen Smith. 2014. *The Epicurean Inscription of Diogenes of Oinoanda: Ten Years of New Discoveries and Research.* Bonn: Habelt.

Hammerstaedt, Jürgen, and Martin Fergusen Smith. 2016. New Research at Oinoanda and a New Fragment of the Epicurean Diogenes (NF 213). *Epigraphica Anatolica* 49. 109–125.

Indelli, Giovanni. 1978. *Polistrato: Sul disprezzo irrazionale delle opinioni popolari.* Naples: Bibliopolis.

Indelli, Giovanni, and Voula Tsouna-McKirahan. 1995. *Philodemus: On Choices and Avoidances*. Naples: Bibliopolis.

Indelli, Giovanni, and Francesca Longo Auricchio. 2020. Hapax legomena in Filodemo. *Cronache Ercolanesi* 50. 119–132.

Indelli, Giovanni, and Francesca Longo Auricchio. 2021. Hapax legomena in Filodemo (2). *Cronache Ercolanesi* 51. 93–107.

— 2022. Hapax legomena in Filodemo (3). *Cronache Ercolanesi* 52. 323–329.

— 2023. Hapax legomena in Filodemo (4). *Cronache Ercolanesi* 53. 185–192.

— 2024. Hapax legomena in Filodemo (5). *Cronache Ercolanesi* 54. 219–226.

— 2025. Hapax legomena in Filodemo (6). *Cronache Ercolanesi* 55. 123–134.

Jensen, Christian. 1907. *Philodemi ΠΕΡΙ ΟΙΚΟΝΟΜΙΑΣ qui dicitur libellus.* Berlin: Weidmann.

Körte, Alfred. 1890. *Metrodori Epicurei Fragmenta*. Leipzig: Teubner.

Laporte. Éric. 2018. Choosing Features for Classifying Multiword Expressions. In Manfred Sailer and Stella Markantonatou (eds.), *Multiword Expressions: Insights from a Multilingual Perspective*, 143–186. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.1182597

Lavorante, Alessia. 2024. La prosa filosofica in forma dialogica: Epicuro, Sul tempo (*PHerc*. 1413/1416). *Cronache Ercolanesi* 54. 29–39.

Leiwo, Martti. 1994. *Neapolitana: A Study of Population and Language in Graeco-Roman Naples.* Helsinki: Societas Scientiarum Fennica.

Long, Anthony A. 1986. *Hellenistic Philosophy: Stoics, Epicureans, Sceptics*, 2nd edn. Berkeley, CA: University of California.

Long, Anthony A., and David Sedley. 1987. *The Hellenistic Philosophers* 1–2. Cambridge: Cambridge University Press.

Long Auricchio, Francesca. 1977. Φιλοδήμου Περὶ ῥητορικῆς *Libros primum* et *alterum.* Naples: Giannini.

— 1985. *Testimonianze dalla Retorica* di Filodemo sulla concezione dell'oratoria nei primi maestri epicurei. *Cronache Ercolanesi* 15. 31–61.

— 1988. *Ermarco: Frammenti.* Naples: Bibliopolis.

Longo Auricchio, Francesca, Giovanni Indelli, Giuliana Leone, and Gianluca Del Mastro. 2020. *La Villa dei Papiri: Una residenza antica a la sua biblioteca.* Rome: Carocci.

McOsker, Michael. 2017. Hiatus in Epicurean Authors. *Cronache Ercolanesi* 47. 145–161.

— 2022. Philodemus of Gadara's Rhythmic Prose. *Rheinisches Museum* 165. 9–49.

— In progress. Translation of Longo Auricchio et al. (2020).

Nicolardi, Federica. 2018. *Filodemo, Il primo libro della Retorica*. Naples: Bibliopolis. https://doi.org/10.1525/rh.2021.39.2.234

— 2019. Aspetti e problemi della stratigrafia nei papiri ercolanesi: lo spostamento a catena di sovrapposti e sottoposti. *Cronache Ercolanesi* 49. 191–215.

Nicolardi, Federica, Daniel Delattre, Gianluca Del Mastro, Richard Janko, and Robert Fowler. 2024. The Final Columns of *PHerc.Paris*. 4 Revealed through Virtual Unwrapping. *Cronache Ercolanesi*. 54. 9–27.

Obbink, Dirk. 1996. *Philodemus: On Piety: Part One.* Oxford: Oxford University Press.

Parisi, Antonio. 2011. Le citazioni poetiche nei papiri ercolanesi: Tre citazioni euripidee

nei papiri di Demetrio Lacone. *Cronache Ercolanesi* 41. 37–50.

— 2012a. Correzioni, abbreviazioni e segni nel *PHerc*. 831. In Paul Schubert (ed.), *Actes du 26e Congrès international de papyrologie*, 599–605. Geneva: Droz.

— 2012b. Osservazioni sul lessico del *PHerc*. 831 (Demetrio Lacone, *opus incertum*). *Cronache Ercolanesi* 42. 111–119.

— 2014. Le forme del sapere astronomico nell'epicureismo: Un saggio di lettura di *PHerc*. 831, IX-XI Körte. *Cronache Ercolanesi* 44. 49–64.

— 2015. Citare il maestro: Due menzioni di Epicuro in Demetrio Lacone. *Cronache Ercolanesi* 45. 19–32.

— 2016. Errori e sensazioni: Per una nuova lettura di *PHerc*. 831 (Demetrio Lacone, *Opus incertum*). *Cronache Ercolanesi* 46. 39–54.

— 2017. Laus physiologiae, παιδεία e parenesis: Una proposta di lettura (*PHerc*. 831, VII-XV Körte). *Cronache Ercolanesi* 47. 41–53.

Sanders, Kirk. 1999. Towards a New Edition of *PHerc*. 831. *Cronache Ercolanesi* 29. 17–30.

Sedley, David. 1973. *On Nature*, Book XXVIII. *Cronache Ercolanesi* 3. 5–83.

Schmid, Wolfgang. 1955. Die Netze des Seelenfängers, Zur Jagdmetaphorik im philosophischen Protreptikos des Demetrios Lakon (Pap. Herc. 831). *Parola di Passato* 10. 440–447.

— 1959. Aus der Arbeit an einem ethischen Traktat des Demetrios Lacon. In Mario Understeiner et al. (eds.), *Epicurea in Memorian Hectoris Bignone, Miscellanea Philologica, 179–195.* Geneva: Istituto di filologia classica.

Smith, Martin F. 1993. *Diogenes of Oenoanda: The Epicurean Inscription*. Naples: Bibliopolis.

— 2003. *Supplement to Diogenes of Oinoanda: The Epicurean Inscription.* Naples: Bibliopolis.

— 2022. Diogenes of Oinoanda: News and Notes XIV (2019–2021). *Cronache Ercolanesi* 52. 383–399.

— 2026. *Urbi et Orbi: The Epicurean Inscription and Prescription of Diogenes of Oinoanda*. Rome: TAB edizioni.

Sudhaus, Sigfried. 1892–1896. *Philodemi Volumina Rhetorica*. Leipzig: Teubner.

Tsouna, Voula. 2012. *Philodemus: On Property Management.* Atlanta, GA: Society for Biblical Literature. https://doi.org/10.2307/j.ctt46n47m

Usener, Hermann. 1887. *Epicurea.* Leipzig: Teubner.

Vergara, Claudio. 2025. *[Filodemo:] La Provvidenza.* Naples: Bibliopolis.

3. Technical writing across space and time: The challenges of lemmatisation and part-of-speech tagging in the historiographical and medical corpora of (post-) classical Greek (5th c. BCE to 2nd c. CE)

Victoria Beatrix Fendel and Matthew T. Ireland[1]

1. History versus anamnesis

The (post)classical Greek medical corpora, i.e. writings associated with Hippocrates (fifth century BCE) and Galen (second century CE), and the (post)classical historiographical corpora, i.e. the writings by Thucydides (fifth century BCE), Xenophon (fifth to fourth centuries BCE), and Polybius (second century BCE) here, pose difficulties for automated lemmatisation and part-of-speech tagging due to their technical register. This register, however, makes them particularly interesting for the study of multi-word expression that are precise to the in-group, of medical or historiographical practitioners, and obscure to the outsider, i.e. the general public. What is *anamnesis* to the former is *history* to the latter.

1 Contributor roles according to CRediT: V.B. Fendel was in charge of conceptualization, methodology, investigation, supervision, and writing; M. T. Ireland was in charge of software. All the translations are our own unless indicated otherwise. Passages are quoted such as to be findable in the Sketch Engine corpora and standard editions without following any specific school of thought.

©2026 V.B. Fendel & M.T. Ireland, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.03

What are the historiographical and medical corpora? A corpus is, generally speaking, "a large and principled collection of natural texts" (Biber and Reppen 2015: 1). The writings by Thucydides, Xenophon, and Polybius on the one hand and by Hippocrates and Galen on the other hand are such large collections of texts. Table 3.1 provides a numeric overview of these.

Author	Number of words	Number of lemmata	Mean tokens per lemma	Number of sentences	Mean sentence length (in words)
Thucydides	165,635	6,858	24.15	4,625	35.81
Xenophon	351,372	11,077	31.72	13,286	26.45
Polybius	314,455	9,751	32.25	12,204	25.77
Hippocrates	422,039	10,136	41.64	29,501	14.31
Galen	3,047,920	16,558	184.08	150,986	20.19

Table 3.1 Overview of the number of words, lemmata, and sentences in the sub-corpora.[2]

They are natural texts if we accept that the preservation history has shaped them, possibly in the sense of scribal edits, minor additions or deletions, and losses (Nicolai 2007: 18; Hoffmann 2005: chap. 8). They are the native speakers of a corpus language (Fleischman 2000: 34). They are principled in that they are associated with a specific discipline, and discourse, respectively. Medical practitioners focussed on physiological processes, historiographical authors on sociopolitical paradigms.[3]

Both disciplines were emergent in the fifth century BCE. Medicine was a marketplace where practitioners were trying to attract patients (clients) and their language needed to be generally comprehensible, rather than highly technical, to drive, rather than prevent, sales (Lateiner 1986: 16; Willi 2003: 95). Historiography was a classroom where "an author—assuming the persona of 'inquirer'—took responsibility himself for giving an account of human affairs in the past" (Schepens 2007: 40) therein instructing and persuading a wide audience which required generally comprehensible, rather than highly technical, language (Nicolai 2007: 15 and 22; Schepens 2007: 42–43).

2 Statistics are drawn from the .vert files of the sub-corpora. Note that the lemma counts are calculated onto the one-to-many correspondence in the lemma column of the .vert file. Script: https://github.com/MatthewIreland/xml_lemmatiser_tagger#. The script counts the number of sentences in each text, the number of distinct lemmata, and the number of times each lemma appears in the text (sorted by frequency of occurrence).

3 Our text selection takes the group of practitioners and the discourses they created, maintained, and fed into as its core criterium. This means that e.g. among the texts associated with the Hippocratic doctors, there are philosophical writings (*De Placitis Hippocratis et Platonis*) and rhetorical ones (*De aere acquis locis*, see also Jouanna (1984: 28)). We view these as created in the same discursive and group settings as the writings that are more narrowly focussed on physiological processes—doctors need ethics (and philosophy) and communication skills (and rhetoric) to perform their duties.

How did the historiographical and medical discourses emerge? Those working to shape the two emergent disciplines shared the approach of obtaining knowledge from empirical evidence (Lateiner 1986: 15; Schepens 2007: 48) and thus two discourses emerged. These are, in Michel Foucault's sense, "complex networks of statements that make knowledge possible; that delimit what can be said, or understood, within a particular discourse; and that determine who can speak (or at least speak with authority or be heard) within that discourse" (Lynch 2014: 121–122). Discourses emerge from groups of people that share experience and/or knowledge through interaction and/or practice, so-called communities of practice (Wenger 1999; Unwin, Hughes, and Jewson 2007). Such communities of practice can be externally recognisable, for example, schools by means of a location and possibly a type of dress code, but visibility can also be limited to the insider perspective, for example, people sharing a specific interest who may only come together for events or in private social-media groups, to use a modern example.

Crucially, "[d]iscourses can arise, be transformed, and disappear; they are fragmentary and incomplete" (Lynch 2014: 120–121). A strong case to this effect has been made for the Hippocratic corpus: "the CH [sc. *Corpus Hippocraticum*] did not produce a Hippocratic discourse; a discourse (Hippocratic or not is another matter) produced the Corpus. Changes in the Corpus are manifestations of how this discourse changed and evolved over time" (Nelson 2016: 126; see also Eijk 2016: 43 who wants to dismember the *Corpus Hippocraticum*; more generally Castro 2025: 317).[4] The Hippocratic medical texts may have generated fierce debate in this regard as they would remain in use for about 2,000 years (Trenery 2020; Jouanna 2012: chap. 16; Paavilainen 2009: chap. 2). Conversely, with the rise of Christianity, historiography's discourse fundamentally changed as "history became a part of a vision of the world and the destiny of man" (Nicolai 2007: 17).

Why are the medical and historiographical corpora so varied? Discourses may give rise to subject-specific terminology, groups may give rise to group-internal terminology, and items do not have to be monovalent. Due to the size of the corpora in question, Sheinfux et al.'s (2019: 41) observation pertains that in very large corpora even elements that otherwise seem fixed can show variation, innovation, and creativity. The size of the corpora further allows to take into consideration the co-text, i.e. the textual surroundings of an element and their cohesion (Lyons 1995: 271; Crystal 2008: 109), the context, i.e. the discursive surroundings of an element and their coherence (Lyons 1995: 264 and 271; Crystal 2008: 119), and "the dynamics between default values and their discursive readjustments as well

4 Part of such a discourse-driven corpus assemblage is that reworkings (cf. Roselli 2016) are a natural part of it. The starting point is the community of practice, the glue that holds it together is the shared knowledge and experience, and the visible trace is a corpus that may not have been perceived as one at all times.

as participants' management of social relationships and identities" (Leitner and Jucker 2021: 705). Essentially, we are not only interested in the referential but also the social meanings and functions of elements (cf. Bentein 2019; Vieregge 2025).

The chapter addresses the following three research questions with regard to (post)classical Greek medical and historiographical corpora:

1. What is technical writing and what are the characteristics of its language?
2. How do technical writings pose difficulties for lemmatisation and part-of-speech tagging?
3. Why are multi-word expressions in technical writings instructive?

The chapter takes a corpus-based approach. Sketch Engine corpora for the relevant sub-corpora are built by means of reuse and modification of previously developed tools. These corpora are used to contextualise previous observations, put forward new observations, and propose avenues for the further investigation of less visible technical corpora in line with the FAIR data principles.

After this introductory *anamnesis*, the chapter falls into five subsections. Section 2 introduces the reader to the corpora and tools. Section 3 addresses research question 1, Section 4 research question 2, and Section 5 research question 3. Section 6 brings together the results and offers conclusions and suggestions for future work.

2. Adopt and adapt

In order to match visibility, accessibility, and searchability of the five sub-corpora—the writings by Thucydides, Xenophon, and Polybius as well as Hippocrates and Galen—, Sketch Engine corpora were built by means of reuse and modification of previously developed tools.

How do the historiographical and medical writings compare? The medical and historiographical corpora are comparable in terms of register, in that texts were written to be read,[5] rather than, for example, performed like plays or speeches, and in that there was a need for comprehensible, rather than highly technical, language, in the fifth-century BCE beginnings of the two disciplines (Section 1, see further Section 3). The corpora are also comparable in terms of discourse, in that the texts reflect their authors' approach of obtaining knowledge from empirical evidence (Section 1). Historiography is about "creating paradigms, predominantly politico-militaristic or ethical ones" (Nicolai 2007: 14; Rood

5 The readership could be as small as the practitioner himself, e.g. personal notes (cf. Langholf 1990: 184–185). The readership could have any number of intentions with the material they read, including the re-working and re-combining of passages (cf. Roselli 2016).

2014: 92; Nicholson 2021) with the historiographer taking on "the persona of 'inquirer'" (Schepens 2007: 40), who has a keen interest in oral and visual testimony as primary sources (Schepens 2007: 47 on Thucydides; Rood 2014: 64 on Xenophon; Schepens 2007: 51 on Polybius). Medicine is about understanding "the human body, its health and diseases" (Lateiner 1986: 15). The physician does an *anamnesis* and observes symptoms in order to treat the patient. Hippocrates' writings were oftentimes so "compact and concise" that they would "require explanation to be understood correctly by the lay reader or beginner" (Boudon-Millot 2018: 293). Galen did this in his commentary work. When Galen diagnoses and treats patients, he "not only gets things right; he can explain how it is that he does so" (Hankinson 2008: 9). Furthermore, the corpora match as regards their diachronic potential, with the beginnings of both disciplines in the fifth century BCE and their development throughout the Hellenistic and Roman periods (Lateiner 1986: 16–17; Nicolai 2007: 16; Schepens 2007: 40), and as regards their potential for multi-word-expression research. Pilot studies on all sub-corpora have returned valuable results (for example, Fendel 2025c; Fendel Submitted-a; Rusten 2020; Fendel 2025d) (see further Section 5). This is why the five sub-corpora have been selected, as despite their differing subject matter, they show a range of similarities ideationally and conceptually (see further Section 4).

The medical and historiographical corpora differ practically in terms of visibility. The medical texts long remained group-internal which was reflected also in the Latinate translation of terms for references to anatomy, diseases, and procedures, for example, German *Keratitis* as opposed to *Hornhautentzündung* "keratitis" (Doody, Föllinger, and Taub 2012; Durling 1993 on Galen; Karimullah 2020; Maclean 2021 on Hippocratic reception; Bouras-Vallianatos and Zipser 2019 on Galenic reception; Hunt 1999 on translation; Montalt and García-Izquierdo 2025; Kalverkämper 2005 on layperson vs. professional since antiquity; Rochette 2005 on translating technical terms in literature; Fögen 2005); the historiographical texts became schoolbook models already in antiquity (Nicolai 2007: 20–22). This situation left traces in the physical and digital editing of relevant texts. Furthermore, the medical and historiographical corpora differ as regards the tailoring of the tagging tool. Thucydides' and Xenophon's literary classical Attic was part of the training data for the originally developed tool, whereas Hippocrates' Ionic was not (Fendel and Ireland 2023).[6] We change this situation by producing Sketch Engine corpora of all the sub-corpora and adapting the tagging tool to the technical language of the corpora.

How can we reuse a tool developed for literary classical Attic historiography, oratory, and prose? The tool developed to tokenise, lemmatise, part-of-speech and

6 There are papyrological texts (e.g. Reggiani (2019), https://medicaltraditions.org/collection/databases/greek) but these come with a new set of issues already at the input-format stage.

morphologically tag a corpus of classical literary Attic historiography, oratory, and prose operates on TEI .xml files as its input format. It relies on whitespaces for tokenisation[7] and the *Perseus Digital Library* for lemmatisation, part-of-speech and morphological tagging. The tool takes a context-free approach and allows for one-to-many correspondences for tags. The tool outputs data in the form of a vertical table in .vert format along with an error.txt file which provides a log of all items for which lemmatisation, part-of-speech or morphological tagging failed. The output files can be implemented into Sketch Engine by means of a modified corpus configuration file (cf. Fendel and Ireland 2023). By means of inspection of the error.txt files, the tool was fine-tuned for classical literary Attic historiography, oratory, and prose and the TEI .xml input format. In fact, Thucydides' and Xenophon' writings were part of the training data.

We expected texts written in a different dialect (for example, Ionic in Hippocrates and some Galen), at a different time (for example, writings by Polybius and Galen), and/or belonging to a different genre and/or register (for example, medical texts) to cause issues. Technologically, the main adaptation related to handling the input format of the texts. The tool was built to operate on TEI .xml files but the digital version of the medical corpora was encoded in plain .xml.[8] Secondly, the tool's functionality to output error.txt files after tagging a text was used to assess recurring issues to do with the digital encoding of the texts. For instance, we noticed issues with the encoding of hyphenation and missing whitespaces posing issues at the tokenisation stage and issues with punctuation being fused to the word end posing issues at the lemmatisation and part-of-speech-tagging stage. Finally, the tool draws for lemmatisation, part-of-speech-tagging, and morphological tagging on the *Perseus Digital Library*, such that we expect technical morphology and lexis to cause issues (see further Section 4). The previously developed tool and corpus and the newly adapted tool and corpus conform to the FAIR principles.

Why FAIR data? FAIR data limits the risk of containing texts within communities of practice and thus sentencing them to a *Rapunzel* sleep. FAIR data is findable, accessible, interoperable, and reusable (Wilkinson et al. 2016). The Sketch Engine corpora for the medical and historiographical texts and the tools to build them described in this chapter are findable, in that a unique and stable identifier (DOI) has been assigned, accessible, in that all

7 This process comes with known errors regarding crasis phenomena (e.g. καὶ ἐάν *kai ean* to κἄν *kan* "even if"). Clitics that are only prosodically merged with the host are unproblematic (e.g. τε *te* "and"). Clitics that are merged with the host morphologically, e.g. Latin *-que* "and", *-ve* "or", and *-ne* (question particle), generate errors but are rare in Greek. These situations are areas for future work. All the errors appear in the error.txt file such that their frequency can be tracked.

8 Tagging script for unstructured text here: https://github.com/MatthewIreland/xml_lemmatiser_tagger/blob/main/scripts/tag_raw.py

corpus files are provided in open access, interoperable, in that rich metadata is provided with the corpus files in the form of a README file[9] and in that the corpus tools are provided open access via Github,[10] and reusable, in that the corpus files and tools are under a CC BY-NC licence which "enables reusers to distribute, remix, adapt, and build upon the material in any medium or format".[11] Furthermore, we subscribe to the concept of Linked Open Data thus contributing to the Web of Data, that is "a global network of interconnected machine-readable information, as opposed to a vast collection of unconnected datasets" (European Union; see also de Santis and Rossi 2019). In future, this will mean that output data files will be renamed with URIs and reformatted from .vert format into JSON or XML.

3. The language of technical writing

We need to define what texts qualify as technical writing before asking what makes the language of technical writing distinctive and resolving the in-group paradox that is attached to any instance of technical writing.

How do we identify technical writing? Most readers would identify the passages in (1) to (3) as technical in nature. The question is why.

(1) "After you complete your initial optimization work, you will be able to more accurately assess the cost of optimization for your **keyword** list. This is a good time to drill down into **long-tail keywords** to look for the most efficient opportunities. Depending on the size and complexity of the **SEO** project, this phase of **keyword** research can be the most labor-intensive." (Enge, Spencer, and Stricchiola 2023: 146, our emphasis)

(2) "Monitor **KPIs** and drive continuous improvement.

Proficiency with **CRM** systems (e.g., Salesforce, ZenDesk) and ticketing platforms is an advantage.

Serve as the quality lead, reviewing and approving all content in the **CMS** workflow."

(from LinkedIn adverts in the areas of customer service and content development, our emphasis)

(3) Hippocrates, *Diseases*, I.4

Τάδε ἐπὶ κακοῖσιν ἀνάγκη κακὰ γίνεσθαι· **ῥῖγος** ἢν λάβῃ, **πῦρ** ἐπιλαβεῖν· καὶ **νεῦρον** ἢν διακοπῇ μὴ ξυμφῦναι **ἐπιφλεγμῆναί** τε ἰσχυρῶς· καὶ ἢν ὁ **ἐγκέφαλος** σεισθῇ τε

9 https://doi.org/10.5281/zenodo.18440945 and https://doi.org/10.5281/zenodo.18441284

10 See https://github.com/MatthewIreland/xml_lemmatiser_tagger (scripts/tag.py and scripts/tag_raw.py).

11 See here https://creativecommons.org/share-your-work/cclicenses/

καὶ πονήσῃ πληγέντος, ἄφωνον παραχρῆμα ἀνάγκη γενέσθαι, καὶ μήτε ὁρῆν μήτε ἀκούειν· ἢν δὲ τρωθῇ, **πυρετόν** τ᾽ ἐπιγενέσθαι καὶ **χολῆς ἔμετον**, καὶ **ἀπόπληκτόν** τι τοῦ σώματος γενέσθαι, καὶ ἀπολέσθαι· **ἐπιπλοῖον** ἢν ἐκπέσῃ, ἀνάγκη τοῦτο **ἀποσαπῆναι**· κἢν αἷμα ἐκ τρώματος ἢ **φλεβὸς** ῥυῇ ἐς τὴν ἄνω **κοιλίην**, ἀνάγκη τοῦτο **πύον** γενέσθαι.

tade epi kakoisin anagkē kaka ginesthai: ***hrigos*** *ēn labē,* ***pur*** *epilabein: kai* ***neuron*** *ēn diakopē mē xumphunai* ***epiphlegmēnai*** *te iskhurōs: kai ēn ho* ***egkephalos*** *seisthē te kai ponēsē plēgentos, aphōnon parakhrēma anagkē genesthai, kai mēte horēn mēte akouein: ēn de trōthē,* ***pureton*** *t᾽ epigenesthai kai* ***kholēs emeton****, kai* ***apoplēkton*** *ti tou sōmatos genesthai, kai apolesthai:* ***epiploion*** *ēn ekpesē, anagkē touto* ***aposapēnai****: kēn aima ek trōmatos ē* ***phlebos*** *hruē es tēn anō* ***koiliēn****, anagkē touto* ***puon*** *genesthai.*

"The following evils follow inevitably upon one another: if there is a **chill**, **fever** follows. If a **cord** is severed, it does not reunite, and a violent **swelling supervenes**. If the **brain** is shaken and suffers damage as the result of a blow, the patient immediately loses his speech, sight, and hearing; if the brain is wounded, **fever** and the **vomiting of bile** ensue, the patient becomes **paralysed** in some part of his body, and he dies. If a **fold of peritoneum** becomes exposed, it must **putrefy**. If blood flows from a wound or a **vessel** into the upper **cavity**, it must turn to **pus**." (Translation by Potter (1988), our emphasis)

The three extracts share at least three features. Firstly, they are highly specific regarding their subject matter. Secondly, they presuppose a background in the subject matter in the reader. Thirdly, they contain lexical items that are limited to the specific subject area.

Observations one and two were made about medical language by Langslow (2000: 26–27), i.e. there is a contextual and a social parameter to the definition of medical language. However, the use of an item by a member of a specific group or community, such as physicians, would not qualify the item as technical as this use may be in an everyday context, like a supermarket (Willi 2003: 52). We thus land on "topic as the basic parameter of technical language" (Willi 2003: 52). The group of practitioners however shapes the topic and discourse (Nelson 2016 on the Hippocratic discourse; Schepens 2007 on the persona of the inquirer; Nicolai 2007 on paradigm presentation). Physicians talk about "the human body, its health and diseases" (Lateiner 1986: 15) with the goal being diagnosis and treatment (if possible); historiographers talk about the "cultural and political achievements of noteworthy communities in the recent past" (Lateiner 1986: 15) with the goal being the creation of paradigms (Nicolai 2007: 14–15).

Conversely, "syntactic or stylistic features that are sometimes listed as 'technical' […] are not technical-language features but register features" (Willi 2003: 53). These features are often applied non-systematically in domain-specific technical writings (Eijk 1997: 101–106 on the Hippocratic corpus) and appear in non-technical contexts or different-domain technical writings too. For example,

aspects of medical language appear in oratory (Mader 2018 on imagery; Das 2019 on similes and conceptual metaphors), historiography (Lateiner 1986; Thomas 2000: chap. 2), tragedy (Clarke Kosak 2004; Jouanna 2012: chaps. 4 and 5), and to a limited extent comedy (Willi 2003: 95). None of these text types deals with medical subject matter.

One such syntactic or stylistic feature deserves brief comment, the so-called nominal style (Panagl 1982: 231; Mambrini 2019; Joho 2022: chap. 1; Ittzés 2024: 93 n. 12; Marušić 2018: 103). The medical texts seem to be innovative in the creation of nominal formations (for example, in -ιτις *-itis* and -ις *-is*) to describe symptoms during *anamnesis*, diseases during *diagnosis*, and anatomical details (Eijk 1997: 101–102). Yet, these formations lack systematicity in formation and exclusivity to the medical domain (Eijk 1997: 102; Willi 2003: 67).

What is distinctive about the language of technical writing? Let us consider a modern example. One author recently took a multi-module course on SEO offered by the University of California Davis. The SEO community is an interesting case study as it is constantly engaging with the user, i.e. all of us, communicating with the user and trying to understand the user, but at the same time it has its toolboxes and tricks of the trade—in fact most of you will have wondered since sentence two in this paragraph what SEO stands for or what it is. The community of practitioners is neither closed off nor secretive as a whole but it thrives on group-internal, subject-specific, and ambiguous terminology (Fendel Submitted-a). A passage from a key manual is in (4) (repeated from above):

> (4) "After you complete your initial optimization work, you will be able to more accurately assess the cost of optimization for your **keyword** list. This is a good time to drill down into **long-tail keywords** to look for the most efficient opportunities. Depending on the size and complexity of the **SEO** project, this phase of **keyword** research can be the most labor-intensive." (Enge, Spencer, and Stricchiola 2023: 146, our emphasis)

A group-internal term is "long-tail keywords". In essence, while keywords such as "laptop" or even "gaming laptop" may be popular and mainstream, they are also highly competitive, i.e. they apply to many different products, companies, services, etc. that one may search for. Conversely, long-tail keywords such as "Lenovo gaming laptop 15 inch" are more specific, usually more niche, and at times even obscure, i.e. you may not think of the search term when you first start searching. A subject-specific term is "search engine optimisation" (SEO), which in fact "is a bit of a misnomer" as SEO is not about optimising search engines themselves but search results (Enge, Spencer, and Stricchiola 2023: 1). An ambiguous term is "keyword", which seems to stand for keyword proper and key phrase in this context. A keyword is "a word (usually one of several) chosen to indicate or represent the content of a larger document, text, record, etc." or a word "entered as a search term in a database or search engine" (Mcarthur,

Lam-Mcarthur, and Fontaine 2018: s.v. keyword). While the original use of the term is anaphoric, the newer use is cataphoric. Furthermore, keywords in SEO are selected not only for their contextual relevancy but also for their competitiveness and form part of the written text rather than functioning as metadata after the fact (see also Wright 2014).

Every text is likely to have ingredients of all three in addition to general-purpose items. For example, if we take a closer look at (3) above from Hippocrates' *Diseases*, *Thesaurus Linguae Graecae* statistics (for the eighth century BCE to third century CE) reveal that only three items in the passage may be subject-specific, as ≥70% of attestations of these appear in the Hippocratic and Galenic writings, see (5).

(5) *Thesaurus Linguae Graecae* statistics (eighth century BCE to third century CE) for Hippocrates, *Diseases*, I.4 (statistics drawn September 2025)

ῥῖγος *rigos* "shivering" (686 of 966 in Galen and Hippocrates) 71%

νεῦρον *neuron* "cord, nerve" (2,545 of 3526 in Galen and Hippocrates) 72%

φλέψ *phleps* "blood-vessel (vein or artery)" (3,163 of 4287 in Galen and Hippocrates) 74%

There is no item that is group-internal, i.e. limited to the in-group of medical practitioners in this extract. There are several ambiguous items, for example, πῦρ *pur* "fire" but also "fever" and κοιλία *koilia* "(any) cavity of the body" but also "cavity (in the earth)".[12] The presence of group-internal and ambiguous items can cause obscurity; the presence of subject-specific items contributes to precision. Medical and historiographical terminology could be considered subject-specific. While medical terms, in their Latinate form, have long dominated in translations of medical writings (Doody, Föllinger, and Taub 2012) and created obscurity to the outside of the community of medical practitioners, historiographical terminology "even for key concepts of method, is largely made up of common words, which may then be used in a less familiar, more technical sense" to this day (Schepens 2007: 43). Thus, while medical terminology seems to be group-internal, historiographical terminology falls into the ambiguous category.

Why is technical writing precise and obscure at the referential and social levels? While we initially may think of group-internal items being very common in technical language, they are never the exclusive component and it crucially depends on the intended audience how much group-internal terminology can be applied in order to avoid compromising comprehensibility (Ferreira 2025: 50–51 on common ground). Consider (2) above. You notice many TLAs (three-

12 While for κοιλία *koilia* the medical meaning seems to have preceded the general meaning, for πῦρ *pur* the general meaning preceded the medical meaning.

letter abbreviations), such as CMS (content management system, for example, WordPress), CRM (a client management system, for example, Salesforce), USPs (unique selling points, of products), and KPIs (key performance indicators, for workers). To the uninitiated these may sound off-putting, initially possibly obscure—although ChatGPT and Google will surely assist you (Sabanés and Cunha 2025). The thus created language variety seems to be of the kind of Blake's argot, i.e. a variety "designed to restrict communication to an in-group and exclude others" (Blake 2010: 227). However, the TLAs are not restricted to an in-group but are used, as in (2), in outward-facing job adverts. They are subject-specific and describe complex workflows, high-pressure environments, team dynamics, and sales targets precisely. This then seems rather akin to Blake's slang, i.e. a variety that is "inventive, abounds in metaphor and allusion, and is often more demanding or more opaque than the standard language" (Blake 2010: 201). New TLAs appear daily, there is no lack of inventiveness.

Inventiveness seems to have been a persistent feature of the medical and historiographical corpora. Polybius's historiographical (second century BCE) and Galen's medical (second century CE) writings illustrate the continuing inventive development of the discourses and their associated language. Polybius famously introduced a range of Latinate technical terms into his Greek and put on a reactionary tone towards the bookish Timaeus (Schepens 2007: 53–54), which shows how fundamental the transformation of Greek historiography at the time was. Galen explained, in the form of commentaries, the "compact and concise" Hippocratic writings referring to them "more than 2,500 times" (Boudon-Millot 2018: 293–294), which reflects the need for updated audience design (see Morison 2008).

While both subject-specific and group-internal terminology is somewhat expected in technical writing, ambiguous items are not. Langslow (2012: 102), for instance, called polysemy and synonymy the "two arch-enemies of technical usage" and Willi (2003: 69) insisted that technical terms must be "standardised in usage and at least ideally monosemous" and "expressively neutral". An interesting case is Polybius' Latinisms, such as στρατηγός *stratēgos* for *consul* but also *praetor*. The *consul* is the highest magistrate in the Roman *cursus honorum*; the *praetor* was one level further down in the *cursus honorum* yet could carry out consular duties in the absence of the *consul*. Langslow (2012: 104) resolves the issue by arguing in favour of audience design, which seems to align with what we see in earlier historians that Polybius likely wanted to emulate (Nicolai 2007: 24).

Furthermore, it may be that the terms are not expressively neutral nor equivalent as regards their indexical value (Eckert 2008; Vieregge 2025). Linguistic indexes are "'structures' (lexemes, affixes, diminutives, syntactic constructions, emphatic stress, etc.) that have become conventionally associated

with a particular situational dimension, and that invoke that situational dimension whenever they are used" (Bentein 2019: 123). Polybius' στρατηγός *stratēgos* may have acquired a conventionalised association with a specific situational dimension and Polybius may have chosen it when he wanted to evoke this situational dimension.

Finally, ambiguous items may function akin to a bridge, not a sturdy one, but one that initially lets the audience feel drawn to the subject matter and community of practice. They may subsequently realise that they were mistaken in their interpretation. This may spark annoyance or curiosity, by the Goldilocks principle (cf. Pierrehumbert 2006: 525; Divjak 2019: 182). Which response it sparks may be a kind of initiation question that these ambiguous items subtly ask.

To resolve the paradox of obscurity and precision simultaneously (Elder 1954), we must consider perspective. What is maximally precise to the in-group, possibly due to group-internal conventions, is likely to be maximally obscure to the outside observer, who would not be privy to group-internal conventions. However, we have noted that in addition to the items that are present for precision (subject-specific) and obscurity (group-internal), there are also bridging items that function somewhat as a gatekeeping phenomenon.[13] Ambiguity is not coincidental but well-planned in order to keep a discourse alive.

4. Lemmatisation and part-of-speech tagging

As outlined in Section 2, we adopted and adapted our tool developed for the tokenisation, lemmatisation, and part-of-speech tagging of classical literary Attic historiography, oratory, and prose in order to convert the raw text of the technical corpora into Sketch Engine corpora which would allow for the further investigation of multi-word expressions. Sketch Engine, as a digital corpus-analysis tool, facilitates the drawing, sorting, and inspecting of collocation tables, for simplex items and combinations of items (lemmata, wordforms, tags) as well as the extraction of n-grams, that is recurringly co-occurring items like δὲ καί *de kai* "but also" in the Hippocratic writings.[14]

Table 3.2 shows the analysis of the errors[15] for the first round of tagging of each sub-corpus. Note that items often incur a lemma and part-of-speech error if they are not listed in the resources the tool draws on. Note further that we have excluded any errors that affected items that begin with a capital letter.

13 The idea of gatekeeping is borrowed from Foucault's (2008) theory of heterotopia (see also Busch and Spitzmüller 2021; Fendel 2026).

14 Currently, no Sketch Grammar for Ancient Greek is implemented such that some functionalities are not accessible.

15 Script here: https://github.com/MatthewIreland/xml_lemmatiser_tagger#. The Python script is run on the .vert files in which errors are recorded in-line. It outputs a table of <word,lemma error count,PoS error count,generic error count,total error count> in CSV format.

As sentence-initial items are not capitalised in the digitised text, items with a capital letter are likely to be personal names. No onomasticon is implemented in the tagging tool currently such that we would expect personal names to cause errors.

Author	Xenophon	Thucydides	Polybius	Hippocrates	Galen
Total words	351,372	165,635	314,455	422,039	3,047,920
Total words with an error	224 <1%	418 <1%	3,564 1%	28,178 7%	127,663 4%
Total lemma errors	100	129	1,219	28,141	127,518
Total part-of-speech errors	99	123	1,215	28,140	127,140
Total generic errors (tech errors)	124	289	2,345	36	137
Total unique errors	203 (of total words with an error:) 91%	397 (of total words with an error:) 95%	2,688 (of total words with an error:) 75%	14,545 (of total words with an error:) 52%	39,152 (of total words with an error:) 31%

Table 3.2 Error analysis from first-round tagging.

A noticeable difference exists between the ratio of total errors, i.e. the total number of tokens that resulted in any type of error during the tagging process, in the Attic historians and Polybius, Hippocrates, and Galen. This is expected given the training data the original tool was based on yet links to a broader ideational issue. Secondly, a noticeable difference exists between the ratio of unique errors, vis-à-vis the total number of errors, in Polybius, Hippocrates, and Galen as compared to the Attic historians. This is expected given the training data the original tool was based on yet also links to a broader conceptual issue. The below unpacks the broader ideational and conceptual issues before suggesting practical solutions.

Ideationally, two long-standing beliefs have made the Attic historians visible and prominent and the later Polybius along with the medical writings comparatively invisible and marginal. The two beliefs are firstly the concept of the so-called classical period, that is the fifth and fourth centuries BCE in Athens, which came to take its place at the heart of curricula (cf. Viti 2024: 20–21; Nicolai 2007: 23), and secondly the concept of literature as implying a “certain aesthetic shaping of a text which is not typically a central concern for authors of modern

specialist and technical works" (Doody, Föllinger, and Taub 2012: 233), who may be more concerned with their writing as a "knowledge-constructing tool" (Montalt and García-Izquierdo 2025). This second belief is closely linked to the idea of topic as a key factor discussed in Section 3. While military and political paradigms were societally suitable and at times even fashionable topics, medical diagnoses and procedures were of interest only to a select few. Both beliefs have yielded to more inclusive views (Totelin 2012 on Galen; Langslow 2012 on Polybius) which have let non-classical and technical writings emerge from the shadows. However, the impact of these beliefs is still felt when it comes to the availability of resources, especially those in digital format, such as textbooks, grammar books, and dictionaries as well as language models and tools.

Conceptually, it is necessary to qualify the notions of *hapax legomena*, group-internal features, and dialectal items. *Hapax legomena* are, traditionally, items that are attested only once in our extant texts. However, Section 3 noted that communities of practice, such as the medical or historiographical ones, develop discourses and as part of this are inventive in their language choices (Blake 2010: 201 on slang). Greek has a fully functioning system of derivational morphology, which is more or less productive, and a fully functioning system of lexical compounding (Fendel 2025e with further references). This means that inventiveness could be implemented in the form of minor formal modifications on an item, even if general systematicity is lacking (Eijk 1997: 102; Willi 2003: 67). This explains the high number of unique errors in Hippocrates and Galen in particular. We may wonder whether such explicable modifications to an item should truly be considered *hapax legomena*.

Group-internal features of any kind, including morphological, syntactic, and lexical ones, are pervasive in corpora produced and managed by the community of practice but do not feature outside of these corpora. They would thus only be considered relevant to include in resources, such as grammar books, textbooks, or dictionaries, which are specifically designed for relevant corpora. This observation is closely linked to the in-group paradox discussed in Section 3. However, fully group-internal features are comparatively rare (Fendel Submitted-a) and relevant features may appear marginally or as one-offs in group-external corpora. This often earns them the label of "exception" or, in the worst case, "mistake", a situation we gradually need to amend across resources as formerly marginal corpora are becoming more visible and accessible and thus allow for comparative work.

Dialectal items are relevant especially to the Hippocratic texts written in the Ionic dialect. Ionic may have functioned as a kind of technolect in the fifth century BCE (Willi 2003: 95). Like any dialect, it contains areal features (here an "East Greek component"), contact features (here "analogies to Attic"), and unique features (García-Ramón and Binder; Colvin 2010: 206–207; Miller

2013: chaps. 13–15; Giannakis 2025). Areal features are unproblematic as the data sample available to examine these features, and train tools on, is relatively large. Furthermore, areal features are likely to be listed in standard dictionaries. For example, Liddell-Scott-Jones lists a "normal or default form", that is a "form [which] could potentially occur in the widest range of Greek texts (with especial weight given to the kinds of texts that are widely read)" (Probert 2019: 209). Contact features are unproblematic as the availability of data and focus of resources is skewed towards Attic, the contact variety in this case. This leaves only Ionicisms proper as a potential challenge. However, Herodotus' *Histories*, written in Ionic, have long been canonical and have thus fed into tools and resources. The situation differs for Polybius' language variety which is diachronically and diastratically distinct from canonical texts (Langslow 2012 on Latinisms; Briscoe 2013 on technical terms; Hutchinson 2025 on book 8).[16] Furthermore, Polybius falls into the timeframe of the post-Koine dialects, that is the local diversification of the Koine due to language-contact scenarios (Bubenik 1993; García Ramón 2020). This explains the high number of unique errors in Polybius.[17]

From the ideational and conceptual issues discussed, two practical issues arise. Firstly, the visibility of texts in digital form as training data for tools (such as the Universal Dependencies models) is limited outside of classical literary Attic.[18] Secondly, the coverage of later and medical texts in open-access resources, including dictionaries that we rely on for lemmatisation, is limited.[19]

A practical solution to the latter issue would be to draw on specialised dictionaries akin to Henderson's (1991) *Maculate Muse* on obscene language in Aristophanic comedy. Some such dictionaries exist for the medical and historiographic corpora, yet often still in formats that are not easily machine readable, for example, Durling collated 3,000 items from the Galenic corpora referring to the fields of anatomy, physiology, pathology, pharmacy, and surgery (Durling 1993: VII). However, he admits to have "excluded all those writings which have been rejected by modern scholarship" (Durling 1993: VIII), which is problematic if we approach the Galenic writings as the work by a community of practice developing a discourse rather than one man's work (cf. Nelson 2016 on Hippocrates), as well as to have shortened lengthier

16 Given the type of corpus we are building, we have focussed on form rather than form-function pairing. It is likely that formally identical items differ in their semantic and pragmatic profiles between dialects as well as diachronically and diastratically.

17 Future work could involve small language models (small LMs), that is engines specialised for specific varieties (see e.g. Tzanoulinou, Triantafyllopoulos, and Verykios 2025). The Daidalos Project at the Humboldt University, Berlin is currently working into this direction for Latin.

18 Biases in training data are well known for modern-language LLMs (Ferreira 2025: chap. 9) and arise equally from canonisation of texts for corpus languages.

19 Older versions of Liddell-Scott-Jones are available open access: https://lsj.gr/wiki/LSJ:GreekEnglishLexicon

entries from the letter E onwards due to "the sheer quantity of material" (Durling 1993: VIII).[20] Nonetheless, the resources exist, even if currently in the wrong format.[21]

The issue concerning the availability of non-canonical texts in digital formats that allow for manipulation and development of tools could be addressed by the large-scale digitisation of relevant corpora, yet this is time-consuming and will initially render few highlight results.[22] More realistic seems the incremental building of smaller corpora, such as the Sketch Engine corpus for the fourth-century CE Kellis papyri (Fendel 2026) or the funerary inscriptions from Roman Sicily (Fendel 2024a). The Sketch Engine corpora, and tools to build them, for the Galenic and Hippocratic texts as well as Polybius' *Histories* are a step into this direction. These will eventually grow into a collection that enhances the visibility of non-canonical, and possibly marginalised, texts.

Specifically regarding multi-word expressions, there is the additional issue that the Universal Dependencies models for (post)classical Greek for instance currently do not operate with a language-specific xpos tag for multi-word expressions. However, efforts to enhance the treebanks and build corpora suitable for the study of multi-word expressions are ongoing, for instance in the form of the PARSEME Ancient Greek corpus (Fendel, Squeri, and Platanou 2026).[23] Furthermore, a workable step forward may be resources like Temple's *Dictionary of Classical Greek Idiom and Syntax* (post-Homeric), which is open access and aims to make phraseological units more visible than they are in standard dictionaries, to learners, translators, and machines.[24]

5. Verbal multi-word expressions

Using multi-word expressions as the structure of interest, we put our Sketch Engine corpora to the test as regards the visibility (and accessibility) of texts—with κενεαγγεία *keneaggeia* "lowering/evacuant treatment, evacuation by bleeding"—and the availability (and usability) of analysis tools—with support-verb constructions of the type verb-object (cf. Mel'čuk 1996; Tronci 2017) with

20 Promising, as they are out of copyright, are the nineteenth-century works on Thucydides and Polybius (*Lexicon Thucydidaeum* (Anonymous 1824) and *Lexicon Polybianum* (Ernesti 1822)), yet these must be considered in the context of their time.

21 It is currently not efficient to annotate manually every lemma that returns NO_LEMMA in the tagging process, yet the error.txt files allow for this and thus to fine-tune the tool to the medical and later historiographic writings.

22 See Ferreira (2025: 96) on publication bias and the Registered Reports community as a counter-initiative.

23 Accessible here https://gitlab.com/parseme/sharedtask-data/-/tree/master/2.0/subtask1. In column 11 of the .cupt file, multi-word expressions are annotated by means of universal annotation guidelines.

24 Accessible here https://sites.google.com/view/classical-greek-idioms/home

πορεία *poreia* “journey”.

A multi-word expression is a multi-lexemic unit that functions as a single unit at the part-of-speech level, for example, the preposition *instead of* or the verb phrase *to make a decision*. Multi-word expressions lend themselves as a test case as they are diachronically, distratically, and diatopically variable in every way, i.e. lexically, morphosyntactically, semantically, and pragmatically (Fendel 2025b; Fendel 2025c). Thus, they provide “sociocultural, economic, and geographical coordinates in time and space” (Tagliamonte 2025: 3). Multi-word expressions are ambiguous, at the lexicon-syntax interface, for example, ὀνόματα κατέχω *onomata katekhō* “to have names” vs. (in GEMF 30.77–81) “to hold names (written on twigs)” with a multi-word-expression as opposed to a verb-object reading (Fendel 2025a: 124), but also semantico-pragmatically, for example, ἐν ἀπόρῳ ἔχομαι *en aporō ekhomai* “to be clueless” is used exclusively in historiography (Thucydides and later only in Lucian and Cassius Dio) and may thus signpost to the community of practice (Fendel 2025d: 103–104).

In the medical corpora, the noun κενεαγγεία *keneaggeia* “lowering/evacuant treatment, evacuation by bleeding” has a subject-specific meaning and seems to be a group-internal term, or even coinage. The related verb κενεαγγέω *keneaggeō* “to have the vessels of the body empty, undergo lowering/evacuant treatment” and the noun in its Ionic form κενεαγγίη *keneaggiē* “lowering/evacuant treatment, evacuation by bleeding” (without iotacism and with raised /a/) appear in Liddell-Scott-Jones as does κεναγγία *kenaggia* “emptiness of vessels” which however seems to differ in usage in that it refers to being without food, in Plato, *Fragments*, 156 “hunger” and in Aristophanes, *Fragments*, 608 “to fast” with ἄγειν *agein* “to lead”.[25] This latter item appears also in the *Perseus Digital Library* tools that underlie the tagging process. κενεαγγεία *keneaggeia* is noticeable because (i) it is not listed (in this form) in standard dictionaries, (ii) it is a group-internal item (having the *Thesaurus Linguae Graecae* Text Search tool operate on the wordform κενεαγγεί *keneaggei* returns hits from medical texts only), and (iii) it is an item that is frequent in medical texts, not a *hapax legomenon* or an innovative coinage.

A string search in the Sketch Engine corpora for the Galenic and Hippocratic writings for κενεαγγε *keneagge* returns 17 hits in the Hippocratic and 27 in the Galenic corpora.[26] String searches are necessary as the lemma, as mentioned,

25 Chantraine (1933: 87–90, §§ 67–68) on the suffixes -ια *-eia* / -ία *-ia* / -εῖᾱ *-eiā* / -εία *-eiē* / -ίᾱ *-iā* / -ίη *-iē* as well as -ηίη *-ēiē* / εῖᾱ *-eiā*. κενεαγγεία *keneaggeia* seems to be a non-Ionic form.

26 The complete annotated dataset is accessible here: https://doi.org/10.5287/ora-jbavbzvz0. As we were interested in whether the item is subject-specific in meaning, it was relevant to inspect the attestations for text types. The lemma appears in the Hippocratic writings only in medical texts *strictu sensu*, i.e. not in rhetorical or philosophical writings, and in the Galenic writings almost exclusively in commentaries on Hippocratic writings.

is not integrated in the tools underlying the tagging process (thus instances are lemmatised as NO_LEMMA). In the 17 Hippocratic passages, the lemma appears seven times in the function of a predicative noun in a verbal multi-word expression, primarily in combination with the verb εἶναι *einai* "to be".[27] In the 27 Galenic passages, the lemma appears twelve times in the function of a predicative noun in a verbal multi-word expression, primarily in combination with the verb εἶναι *einai* "to be". The Sketch Engine collocations reveal that the lemma is not a cranberry item, i.e. "a token that does not have the status of a stand-alone word, has no proper distribution, and no stand-alone meaning, but it may have a syntactic category and an inflection paradigm. It only occurs in a particular expression (or a closed list of expressions) and can never be found in different contexts".[28] Thus, while κενεαγγεία *keneaggeia* may be group-internal in usage and subject-specific in meaning, it is not a *hapax legomenon*, cranberry item, or innovative coinage. Nonetheless, it lacks visibility in standard dictionaries and tools. For future iterations, analysis of the error.txt files allows us to identify and manually (if necessary) tag items like this lemma and train the tool to recognise these. As the number of Hippocratic and Galenic texts searchable by means of our Sketch Engine corpora is significantly higher than the number of those searchable in the *Thesaurus Linguae Graecae*, we are better equipped to avoid misidentification of items as *hapax legomena*, cranberry items, or mistakes.

In Polybius' *Histories*, journeying is central (Rood 2012). It is referred to considerably less often by means of a simplex verb, πορεύομαι *poreuomai* "to journey", than a verbal multi-word expression, with the noun πορεία *poreia* "journey" (Fendel 2025c: 529). With regard to the latter, a subject-specific coinage seems to arise in Polybius' writings that comes to signpost to the specific text type and community of practice producing it.

Of the 183 hits for the lemma πορεία *poreia* "journey" in Polybius' *Histories*, the lemma functions as a predicative noun in a verbal multi-word expression in 121. Thus, it is not a cranberry item. Table 3.3 provides an overview of the lexical and morphosyntactic characteristics of these 121 instances. All the instances show πορεία *poreia* "journey" as the predicative noun in a verb-noun-type verbal multi-word expression, which we term a support-verb construction with the verb being a support verb. By contrast, the simplex lemma πορεύομαι *poreuomai* "to journey" (always in the middle voice) appears only 34 times in Polybius' *Histories*.

27 On εἰμί *eimi* as a support verb, see Fendel (Submitted-c).

28 PARSEME Universal Annotation guidelines for multi-word expressions 2.0: https://parsemefr.lis-lab.fr/parseme-st-guidelines/2.0/index.php?page=glossary#cranberry-word

Support verb/Verb of realisation	Number of the predicative noun	Determiner phrase	Attributive phrase	Word order support verb-predicative noun	Mean support-verb-construction field	Total
ποιέω *poieō* "to do"	Singular	100%	15%	99%	0.25	95 79%
	Plural	100%	0%	67%	0	3 2%
χράομαι *khraomai* "to use"	Singular	89%	56%	78%	0.22	9 7%
	Plural	100%	100%	100%	0.5	2 2%
Verb of realisation	Singular	100%	50%	78%	1.25	10 8%
	Plural	50%	50%	100%	0	2 2%

Table 3.3 πορεία *poreia* "journey" in Polybius' *Histories*.

There are two main support verbs with πορεία *poreia* "journey", that is ποιέω *poieō* "to do" and χράομαι *khraomai* "to use", along with a range of largely one-off (except for ἄγω *agō* "to lead" which appears twice) verbs of realisation. Verbs of realisation are "des verbes collocationnels qui ont le comportement syntaxique des Vsupp, mais qui, à la différence de ceux-ci, sont sémantiquement pleins" (translation: "collocating verbs which behave syntactically like Vsupp but, unlike them, are semantically full/heavy") (Mel'čuk 2004: 208).

Generally, there is a strong preference for the singular form of the predicative noun. This may be linked to the fact that the plural of the predicative noun would be particularising whereas the singular seems to be treated like as a mass noun (Wittenberg and Levy 2017; Erbach 2025). The combination with ποιέω *poieō* "to do" tends towards invariability, especially morphosyntactically (for example, a determiner phrase is always present) and syntactically (i.e. the support verb-predicative noun order appears in 99% of cases and only a mean of 0.25 items intervene between constituents). Where word-order changes, an attributive phrase, or the lack of a determiner phrase were contextually required, the alternative with χράομαι *khraomai* "to use" seems to have been used (see similarly Fendel Submitted-b). If flexibility with regard to the support-verb-construction field was required additionally and thus usually also an additional

semantic or pragmatic nuance, a verb of realisation was used instead.[29] There is no discernible semantic difference between the option with ποιέω *poieō* "to do" and that with χράομαι *khraomai* "to use", which is specifically obvious in Polybius, *Histories*, 8.28.13 and 10.49.3 where the support-verb construction appears in the form of a construct participle with a verb of movement as the finite verb (ἦκον *hēkon* "they came", προῆγε *proēge* "they advanced").[30]

In the classical writings by Thucydides and Xenophon, there are 59 instances of πορεία *poreia* "journey" in total, but only in 15 of these the lemma functions as a predicative noun. All the relevant hits come from Xenophon. Table 3.4 provides an overview of the lexical and morphosyntactic characteristics of these 15 instances. By contrast, the simplex lemma πορεύομαι *poreuomai* "to journey" (always in the middle voice) appears 469 times in Thucydides' and Xenophon's writings.

Support verb/Verb of realisation	Number of the predicative noun	Determiner phrase	Attributive phrase	Word order predicative noun-support verb	Mean support-verb construction field	Total
ποιέω *poieō* "to do"	Singular	100%	14%	86%	0.71	7
	Plural	100%	50%	100%	1	2
γίγνομαι *gignomai* "to happen"	Singular	100%	100%	100%	n/a	1
Verb of realisation	Singular	100%	50%	25%	0	4
	Plural	100%	100%	100%	n/a	1

Table 3.4 πορεία *poreia* "journey" in the writings by (Thucydides and) Xenophon.

There are two main support verbs with πορεία *poreia* "journey", that is ποιέω *poieō* "to do" and γίγνομαι gignomai "to happen" (cf. Jiménez López 2021), along with a range of one-off verbs of realisation. There is overall a strong preference for the singular over the plural form of the predicative noun both with support verbs and with verbs of realisation. The combination with ποιέω *poieō* "to do" tends towards invariability, especially morphosyntactically (for example, a determiner phrase is always present) and syntactically (i.e. the predicative noun-support verb order appears in 86% of cases). Where word-order changes or an attributive phrase

29 There are some combinations where a quantifier or qualifier appears in the object slot of the support verb and the predicative noun is attached as a qualifying genitive. Attributive adjectives if present usually appear outside of the support-verb construction in predicative position and only rarely in attributive position between the determiner phrase and the predicative noun.

30 The complete annotated dataset is accessible here: https://doi.org/10.5287/ora-jbavbzvz0

were contextually required, the alternative with verbs of realisation seems to have been used. χράομαι *khraomai* "to use" does not appear as an option.

There are three noticeable differences between Polybius' as compared to the classical historian's use of the support-verb constructions with πορεία *poreia* "journey". Firstly, the word-order preference changes from predicative noun-support verb in classical times to support verb-predicative noun in post-classical times. This may be due to systemic changes in Greek word order (Horrocks 1990; Wahlgren 2024; Aresu and Cabriolu 2024). Yet the fact that these changes affect the internal order of the support-verb construction suggests that the support-verb construction was analytic enough to allow for this (cf. Fendel 2024b; Fendel 2023). Secondly, there is a numerical preference for the simplex verb over the support-verb construction in classical times which is fully reversed in post-classical times. This may be linked to the systemic change to the middle voice (Gignac 1976: 321–327; Mandēlaras 1973: 369–371; Vives Cuesta and Acero 2022). Yet it may also be that the support-verb construction was advantageous for discourse coherence and cohesion (Fendel and Ireland 2023) or carried an indexical meaning. Thirdly, the support verb χράομαι *khraomai* "to use" does not appear as an option in the classical historians but seems Polybius' go-to option when increased flexibility of form was needed. Table 3.5 shows the distribution of πορείᾳ / (πορείαις) χράομαι *poreia* / (*poreiais*) *khraomai* diachronically in literary texts.

Time period	In historiography	Total
Archaic Greek (pre 5th c. BCE)	–	–
Classical Greek (5th/4th c. BCE)	–	–
Hellenistic Greek (3rd–1st c. BCE)	12 (incl. Polybius and Diodorus Siculus)	12
Roman Greek (1st–3rd c. CE)	3 (in Josephus, Plutarch, Pausanias, Vita Sancti Abercii, Cassius Dio, Diogenes Laertius; note that there are additionally six instances in biographies and hagiography which border onto historiography)	17
Early Byzantine Greek (4th c. – 7th c. CE)	2 (in Zosimus, Procopius)	9
Medieval Greek (post-8th c. CE)	9 (incl. much commentary work and hence quotes; in historiography and hagiography)	19
Total	26	57

Table 3.5 Proximity search for the lemmata χράομαι *khraomai* "to use" and πορεία *poreia* "journey" in the *Thesaurus Linguae Graecae* within a distance of 5 items and manual correction for support-verb constructions.

Table 3.5 shows that while the option with χράομαι *khraomai* "to use" was coined in Polybius, it was subsequently taken up only within the community of practice of historiographers and thus likely acquired an index towards this text type and/or group. In Roman times, the support-verb construction seems to have expanded beyond this initial context, as in (6) from Galen:

(6) Galen, *De usu partium*, Kühn vol. 4 p. 306 l. 15

διὰ μὲν γὰρ τῆς ἔνδον χώρας τοῦ βραχίονος ἐπὶ τὸν πῆχυν ἔρχεται βύθια, πλησίον δ' ἀφικόμενα τῆς κατ' ἀγκῶνα διαρθρώσεως ἀσάρκου καὶ ὀστώδους ὅλης οὔσης ἐκινδύνευσεν ἂν ἐπιπολῆς ὑπὸ τῷ δέρματι γυμνῷ σαρκῶν τοῖς ὀστοῖς ἐποχούμενα **σφαλερωτάτῃ χρήσασθαι πορείᾳ**, εἴπερ μηδὲν ἡ φύσις εὗρεν εἰς ἀσφάλειαν σόφισμα τοιοῦτον, οἷον νῦν ἐστι.

*dia men gar tēs endon khōras tou brakhionos epi ton pēkhun erkhetai buthia, plēsion d' aphikomena tēs kat' agkōna diarthrōseōs asarkou kai ostōdous holēs ousēs ekinduneusen an epipolēs hupo tō dermati gumnō sarkōn tois ostois epokhoumena **sphalerōtatē khrēsasthai poreia** eiper mēden hē phusis heuren eis asphaleian sophisma toiouton, hoion nun esti.*

"For they (sc. nerves) go deep (βύθια *buthia*) through the inner regions of the arm into the forearm, having come near the elbow joint, which is fleshless and made of bone (?) entirely, and are at risk, if they come to the surface under the naked skin on the bones with flesh, **to make a journey that makes one stumble**, if nature had not found for safety such a solution as it is now."

In (6), the support-verb construction in question refers to the movement of nerves. The passage retains the conceptual metaphor of PATH from the original meaning of the support-verb construction. However, the support-verb construction refers no longer to movement across geographical space but to movement more generally, in this case within the body (cf. Jiménez Delgado and Martínez Vázquez 2011). Noticeably, Galen uses the option with χράομαι *khraomai* "to use" when he requires the support-verb construction to be variable enough to allow for an adjective qualifying the predicative noun (σφαλερωτάτῃ *sphalerōtatē* "very likely to make one stumble"). This aligns with Polybius' use and the observation that the option with χράομαι *khraomai* "to use" was created in order to restore flexibility in a support-verb construction that had lost it, i.e. the option with ποιέω *poieō* "to do".[31] While the option with ποιέω *poieō* "to do"

31 For comparison, a proximity search for the lemmata ποιέω *poieō* "to do" and πορεία *poreia* "journey" in the *Thesaurus Linguae Graecae* within a distance of 5 items and manual correction for support-verb constructions returns the following: Archaic Greek – 1 instance, Classical Greek – 37 instances, Hellenistic Greek – 158 instances, Roman Greek – 92 instances, Early Byzantine Greek – 255 instances. Thus, this support-verb construction was not replaced, but the χράομαι *khraomai* "to use" option enhanced the repertoire. The fixedness of structures can be examined by means of inspecting N-grams in Sketch Engine, which was not possible for the full corpus of the *Thesaurus Linguae Graecae*.

is listed in Liddell-Scott-Jones (s.v. πορεία *poreia* II.2), the option with χράομαι *khraomai* "to use" is not. This is partially possibly due to its infrequency in classical corpora, partially due to its infrequency vis-à-vis highly established ποιέω *poieō* "to do", and partially due to the fact that support-verb constructions are not consistently integrated in the dictionary (Fendel 2024c: 328–330). Our Sketch Engine corpora facilitate the investigation of such marginal (subject-specific) structures.

In the first book of Thucydides' *Histories*, a support-verb construction is ambiguous between textual and metatextual meanings, as Rusten (2020) has shown. The relevant passage, (7), was long taken to refer to a digression in the narrative (textual level) but is in fact to be taken as the discarding of a plan (metatextual level):

(7) Thucydides, *Histories*, 1.97.2

ἔγραψα	δὲ	αὐτὰ	καὶ	**τὴν**
egrapsa	*de*	*auta*	*kai*	*tēn*
write.aor.ind.act.1sg	prt	they.acc.pl.n	and	the.acc.sg.f
ἐκβολὴν		**τοῦ**	**λόγου**	
ekbolēn		*tou*	*logou*	
digression/discarding.acc.sg.f		the.gen.sg.m	narrative/plan.gen.sg.m	
ἐποιησάμην	διὰ	τόδε,	ὅτι (...)	
epoiēsamēn	*dia*	*tode*	*hoti*	
do.aor.ind.mid.1sg	due.to	it.acc.sg.n	that	

"I have written these things and (old translation:) made a digression in the narrative / (Rusten's translation:) discarded the plan due to the fact that (...)"

Rusten's re-interpretation rests on the meaning of λόγος *logos* in Thucydides ("content of speech, plan") as opposed to other historiographers (for example, "unit of narrative" in Herodotus). The meaning of the event noun ἐκβολή *ekbolē* "tossing out" is ambiguous between "tossing out for consideration" and "tossing out due to rejection". The co-textually defined meaning of the dependent genitive pushes us towards the latter meaning (Rusten 2020: 234–235). However, noticeably, this metalinguistic remark is rather isolated in Thucydides' writings as he rarely surfaces in his *Histories* as the narrator (Beasley 2023), unlike, for example, Herodotus in his *Histories*.

Even if we adopt Rusten's proposal, the ambiguity may have remained in place for a large proportion of the readership and may have been intentional in

the sense of the bridges mentioned in Section 3. Those readers who were part of the community of practice may have been familiar with Thucydides' specific use of λόγος *logos* and his invisibility as a narrator, those who were at the outskirts of the community or were looking in from the outside may not have had this kind of knowledge. The ambiguity may thus serve as a gatekeeper.

6. Summary, conclusion, and outlook

The chapter has discussed the building and use of Sketch Engine corpora for the (post)classical Greek medical, i.e. writings associated with Hippocrates and Galen, and historiographical corpora, especially the writings by Thucydides, Xenophon, and Polybius, based on the digitised texts of the *Digital Corpus of Graeco-Arabic Studies* and the *Perseus Digital Library*.[32] The medical and historiographical corpora have been considered orderly and principled collections of texts (i.e. corpora) from the perspective of a community of practice maintaining and curating a discourse which is reflected in their writings (Section 1). We adopted and adapted tools previously built and have provided a pipeline that allows for further modification and development (Section 2). The Sketch Engine corpora enhance the visibility of these technical writings, make them accessible and searchable, and offer insights especially from verbal multi-word expressions to enhance existing resources and tools.

We have taken topic as the key criterion to define what texts belong with each of the two technical discourses of interest. The texts of each discourse reflect a mixture of language features, including features that are specific to the in-group (group-internal features), those that are specific to the subject matter discussed (subject-specific features), those that are ambiguous between in-group and generic readings (ambiguous features), and entirely generic features. Furthermore, it appears that technical writings align with Blake's slang in being inventive and at times opaque while at the same time striving for precision. This apparent paradox resolves when we consider that what is maximally precise to the in-group, for example, in the form of group-internal language features, may be maximally obscure to the outsider. Ambiguous features adopt the role of gatekeeping phenomena that can be interpreted correctly by the in-group and may appeal to but eventually deceive the outsider (Section 3).

Such obscurity to the outsider is one factor that has had a far-reaching impact on the visibility and coverage of technical writings in tools and resources. Conceptually, we identified the notion of a classical period and the notion of literature as an aesthetic ideal as critical to technical writings lacking visibility

32 For issues surrounding the textual edition available via these repositories, see Fendel et al. (2026).

for a long time. Ideationally, this resulted in the misidentification of *hapax legomena*, features falsely being labelled exceptions or even mistakes, and features being considered marginal incorrectly. Practically, this explains the issues surrounding the visibility of technical writings in digitised corpora and their coverage by available tools. Our Sketch Engine corpora and the tools and pipelines to build them in line with the FAIR data principles are a significant step towards addressing these issues (Section 4).

Verbal multi-word expressions are a language feature that lacks visibility and coverage in corpora, resources, and tools despite recent advances (for example, Fendel and Ireland 2023; Fendel 2024a; Fendel, Squeri, and Platanou 2026; Fendel 2025c). In the medical and historiographical corpora, κενεαγγεία *keneaggeia* "lowering/evacuant treatment, evacuation by bleeding" illustrates the necessity of comprehensive corpora in order to avoid omission and mislabelling of group-internal and/or subject-specific items in resources and corpora. In the historiographical corpora, subject-specific support-verb constructions of the type verb-object with πορεία *poreia* "journey" illustrate the synchronic and diachronic variability in referential and social meanings of support-verb constructions which can only be uncovered by means of large searchable and suitably annotated corpora, especially when comparatively rare structures are at issue. We should add that currently pragmatic and metatextual annotation is a *desideratum* in our corpora such that instances like Thucydides, *Histories*, 1.97.2 remain difficult to discover. While their analysis by means of comparative evidence is facilitated by the corpora we have built, these gatekeeping devices are still gatekeeping.

Acknowledgements

This work received support from the CA21167 COST action UniDive, funded by COST (European Cooperation in Science and Technology).

Bibliography

Anonymous. 1824. *Lexicon Thucydidæum: A Dictionary in Greek and English of the Words, Phrases, and Principal Idioms Contained in the History of the Peloponnesian War.* London: R. Gilbert.

Aresu, Mauro, and Silvia Cabriolu. 2024. Observations on the Variation of Word-order: A Comparison of Preclassical, Classical and Biblical Latin. In Carlotta Viti (ed.), *Ancient Greek and Latin in the Linguistic Context of the Ancient Mediterranean*, 413–438. Tübingen: Narr.

Beasley, Tom. 2023. Thucydidean Self-Presentation. In Polly Low (ed.), *The Cambridge Companion to Thucydides*, 31–49. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316227442.005

Bentein, Klaas. 2019. Dimensions of Social Meaning in Post-classical Greek: Towards an Integrated Approach. *Journal of Greek Linguistics* 19(2). 119–167. https://doi.org/10.1163/15699846-01902006

Biber, Douglas, and Randi Reppen (eds.). 2015. *The Cambridge Handbook of English Corpus Linguistics.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139764377

Blake, Barry. 2010. *Secret Language*. New York: Oxford University Press.

Boudon-Millot, Véronique. 2018. Galen's Hippocrates. In Peter Pormann (ed.), *The Cambridge Companion to Hippocrates*, 292–314. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781107705784.014

Bouras-Vallianatos, Petros, and Barbara Zipser (eds.). 2019. *Brill's Companion to the Reception of Galen.* Leiden; Boston: Brill. https://doi.org/10.1163/9789004355385

Briscoe, John. 2013. Some Misunderstandings of Polybius in Livy. In Bruce Gibson and Thomas Harrison (eds.), *Polybius and his World: Essays in Memory of F.W. Walbank*, 117–124. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199608409.003.0005

Bubenik, Vit. 1993. Dialect Contact and Koineization: The Case of Hellenistic Greek. *Journal of the Sociology of Language* 99(1). 9–23.

Busch, Brigitta, and Jürgen Spitzmüller. 2021. Indexical Borders: The Sociolinguistic Scales of the Shibboleth. *International Journal of the Sociology of Language* 2021(272). 127–152. https://doi.org/10.1515/ijsl-2020-0095

Castro, Paula. 2025. Possibilities of Existence—Making and Changing Subjectivities and (Ancient) Worlds. In Luuk Huitink, Vlad Glaveanu, and Ineke Sluiter (eds.), *Social Psychology and the Ancient World: Methods and Applications*, 309–318. Leiden; Boston, MA: Brill. https://doi.org/10.1163/9789004731301_017

Chantraine, Pierre. 1933. *La formation des noms en grec ancien.* Paris: É. Champion.

Clarke Kosak, Jennifer. 2004. *Heroic Measures: Hippocratic Medicine in the Making of Euripidean Tragedy.* Leiden; Boston, MA: Brill. https://doi.org/10.1163/9789047405955

Clarysse, Willy. 1993. Egyptian Scribes Writing Greek. *Chronique d'Egypte* 68(135–136). 186–201.

Colvin, Stephen. 2010. Greek Dialects in the Archaic and Classical Ages. In Egbert Bakker (ed.), *A Companion to the Ancient Greek Language*, 200–212. Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9781444317398.ch14

Crystal, David. 2008. *A Dictionary of Linguistics and Phonetics*. 6th edn. Malden, MA; Oxford: Blackwell. https://doi.org/10.1002/9781444302776

Das, Allison E. 2019. Health, Harm, and the Civic Body: Medical Language in the Speeches of Demosthenes. *Greek, Roman, and Byzantine Studies* 59(3). 340–367.

Divjak, Dagmar. 2019. *Frequency in Language: Memory, Attention and Learning*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316084410

Doody, Aude, Sabine Föllinger, and Liba Taub. 2012. Structures and Strategies in Ancient Greek and Roman Technical Writing: An Introduction. *Studies in History and Philosophy of Science* Part A 43(2). 233–236. https://doi.org/10.1016/j.shpsa.2011.12.021

Durling, Richard. 1993. *A Dictionary of Medical Terms in Galen*. Berlin; Boston, MA: Brill.

Eckert, Penelope. 2008. Variation and the Indexical Field. *Journal of Sociolinguistics* 12(4). 453–476. https://doi.org/10.1111/j.1467-9841.2008.00374.x

Eijk, Philip van der. 1997. Towards a Rhetoric of Ancient Scientific Discourse. In Egbert Bakker (ed.), *Grammar as Interpretation: Greek Literature in its Linguistic Contexts*, 77–129. Leiden; Boston, MA: Brill.

— 2016. On 'Hippocratic' and 'Non-Hippocratic' Medical Writings. In Lesley Dean-Jones and Ralph Rosen (eds.), *Ancient Concepts of the Hippocratic: Papers Presented at the XIIIth International Hippocrates Colloquium, Austin, Texas, August 2008*, 15–47. Leiden; Boston, MA: Brill. https://doi.org/10.1163/9789004307407_003

Elder, Joseph. 1954. Jargon—Good and Bad. *Science* 119(3095). 536–538.

Enge, Eric, Stephan Spencer, and Jessie Stricchiola. 2023. *The Art of SEO: Mastering Search Engine Optimization*. 4th edn. Sebastopol, CA: O'Reilly.

Erbach, Kurt. 2025. Shifting Interpretations: Count and Mass in Linguistic Categorization. *Lexis. Journal in English Lexicology* 25. https://doi.org/10.4000/14i18

Ernesti, Johann August. 1822. *Lexicon Polybianum*. Oxford; London: W.B. Whittaker and J. Parker.

European Union. Linking Data: What Does It Mean? (23 October 2025), https://data.europa.eu/en/publications/datastories/linking-data-what-does-it-mean

Evert, Stefan. 2009. Corpora and Collocations. In Anke Lüdeling and Merja Kytö (eds.), *Volume 2: An International Handbook*, 1212–1248. Berlin; Boston, MA: Mouton De Gruyter. https://doi.org/10.1515/9783110213881.2.1212

Fendel, Victoria. 2023. 'I Haven't Got a Clue!': Assessing Negation in Classical Greek Support-Verb Constructions. *Journal of Greek Linguistics* 23(2). 139–163. https://doi.org/10.1163/15699846-02302004

— 2024a. The I.Sicily Sketch Engine Corpus. *Journal of Open Humanities Data* 10. https://doi.org/10.5334/johd.258

— 2024b. Celebrating Diversity: The Origins and Pathways of Three Support-verb Constructions. *Lexis* 24. https://doi.org/10.4000/12cvp

— 2024c. Epilogue: Taking Wing. In Victoria Fendel (ed.), *Support-verb Constructions in the corpora of Greek: Between Lexicon and Grammar?, 327–340*. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.13836407

— 2025a. Support-verb Constructions in the Magical Papyri of the Later Roman and Early Byzantine Periods: A Special Case of Technical Writing. *Symbolae Osloenses* 98(1). 106–136. https://doi.org/10.1080/00397679.2024.2428117

— 2025b. Taking Stock of Greek Support-verb Constructions: Synchronic and Diachronic Variability in the Documentary Papyri. In Jesus de la Villa (ed.), *Advances in Ancient Greek Linguistics*, 295–311. Berlin; Boston, MA: Mouton De Gruyter. https://doi.org/10.1515/9783111648644-017

— 2025c. *Giving Gifts and Doing Favours: Literary Classical Attic Greek Support-verb Constructions.* Leiden; Boston, MA: Brill. https://doi.org/10.1163/9789004730526

— 2025d. Do Good or Suffer Evil: Syntactic Nominalisations in Verbal Multi-Word Expressions in the PARSEME GRC Corpus. *Digital Studies in Language and Literature* 2(1). 85–121. https://doi.org/10.1515/dsll-2025-0005

— 2025e. “To Do or Not to Do”: Semi-lexical Affixes in (Post)classical Greek. *Finnish Journal of Linguistics* 37. 33–72. https://doi.org/10.61197/fjl.142559

— 2026. Le contrôle des accès dans l‘hétérotopie kellienne. In Victoria Fendel, Robin Meyer, and Antoine Viredaz (eds.), *Language and Identity in Antiquity: Graeco-Roman and Ancient Near Eastern Perspectives*. Basel: Schwabe.

— Submitted-a. Die Produktivität von Stützverbgefügen in den medizinischen Texten des (post-)klassischen Griechisch. *Hyperboreus*.

— Submitted-b. Making Conversation and Using Words: Boundedness with the Support Verbs ποιέομαι poieomai “to Do” and χράομαι khraomai “to Use”. *Journal of Greek Linguistics*.

— Submitted-c. From Being, Becoming, and Existing: (Dis)connections between Lexical, Semi-lexical, and Copular Uses of the Being-verb Paradigm. In Ricardo Alcocer Urueta (ed.), *Linguistic Insights into the Meaning and Function of the Verb εἰμί in Classical Greek Philosophy*. Leiden; Berlin: De Gruyter Brill.

Fendel, Victoria, and Matthew Ireland. 2023. Discourse Cohesion in Xenophon’s On Horsemanship through Sketch Engine. *Digital Humanities Quarterly* 17(3). https://doi.org/10.63744/scm43z65uf3d

Fendel, Victoria, Elena Squeri, and Paraskevi Platanou. 2026. Let’s Square a Circle: The PARSEME-GRC Corpus of Verbal Multi-word Expressions. *Journal of Greek Linguistics*.

Ferreira, Fernanda. 2025. *Psycholinguistics: A Very Short Introduction.* Oxford, New York: Oxford University Press. https://doi.org/10.1093/actrade/9780192886774.001.0001

Fleischman, Suzanne. 2000. Methodologies and Ideologies in Historical Linguistics: On Working with Older Languages. In Susan Herring, Pieter Reenen, and Lene Schøsler (eds.), *Textual Parameters in Older Languages*, 33–58. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.195.03fle

Fögen, Thorsten. 2005. Antike Fachtexte als Forschungsgegenstand. In Thorsten Fögen (ed.), *Antike Fachtexts/Ancient Technical Texts*, 1–20. Berlin; New York: de Gruyter. https://doi.org/10.1515/9783110912104.1

Foucault, Michel. 2008. Of Other Spaces* (1976). In Lieven De Cauter and Michiel Dehaene (eds.), *Heterotopia and the City: Public Space in a Postcivil Society*, 13–29. Abingdon: Routledge. https://doi.org/10.4324/9780203089415-8

García Ramón, José Luis. 2020. Change in Grammatical and Lexical Structures in Postclassical Greek: Local Dialects and Supradialectal Tendencies. In Dariya Rafiyenko and Ilja A. Seržant (eds.), *Postclassical Greek: Contemporary Approaches to Philology and Linguistics*, 303–336. Berlin; Boston, MA: Mouton De Gruyter. https://doi.org/10.1515/9783110677522-013

García-Ramón, José Luis, and Vera Binder. Ionic. In *Brill’s New Pauly Online*. Leiden; Boston, MA: Brill. https://doi.org/10.1163/1574-9347_bnp_e526520

Giannakis, Georgios K. 2025. Linguistic Variation and the Study of Ancient Greek Dialects. In Jesús de la Villa, Araceli Striano, and Rodrigo Verano (eds.), *Advances in Ancient Greek Linguistics*, 3–32. Leiden; Berlin: De Gruyter Brill. https://doi.org/10.1515/9783111648644-001

Gignac, Francis. 1976. *A Grammar of the Greek Papyri of the Roman and Byzantine Periods: Morphology.* Milan: Istituto editoriale cisalpino-La goliardica.

Hankinson, Robert. 2008. The Man and his Work. In Robert Hankinson (ed.), *The Cambridge Companion to Galen*, 1–33. Cambridge: Cambridge University Press. https://doi.org/10.1017/ccol9780521819541.001

Henderson, Jeffrey. 1991. *The Maculate Muse: Obscene Language in Attic Comedy*. Oxford: Oxford University Press.

Hoffmann, Sebastian. 2005. *Grammaticalization and English Complex Prepositions: A Corpus-Based Study.* Florence, US: Taylor & Francis. https://doi.org/10.4324/9780203008225

Horrocks, Geoffrey. 1990. Clitics in Greek: A Diachronic Review. In Anna Roussou and Stavros Panteli (eds.), *Greek outside Greece II*, 35–52. Athens: Diaspora Books for the GLRG in collaboration with Interworld Publications.

Hunt, Tony. 1999. Old French Translations of Medical Texts. *Forum for Modern Language Studies* 35(4). 350–357.

Hutchinson, Gregory. 2025. *Polybius: Book 8*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009265843

Ittzés, Máté. 2024. Action Nouns in Vedic Support-verb Constructions with kṛ. *Journal of South Asian Languages and Linguistics* 11(1–2). 87–109. https://doi.org/10.1515/jsall-2024-2008

Jiménez Delgado, José Miguel, and Rafael Martínez Vázquez. 2011. Verbos de movimiento virtual en griego antiguo. *Emerita* 79(2). 277–300. https://doi.org/10.3989/emerita.2011.12.1018

Jiménez López, María Dolores. 2021. Γίγνομαι as the Lexical Passive of the Support Verb ποιέω in Ancient Greek. In Georgios Giannakis, Luz Conti, Jesus de la Villa, and Raquel Fornieles (eds.), *Synchrony and Diachrony of Ancient Greek: Language, Linguistics and Philology*, 227–240. Berlin; Boston, MA: Mouton De Gruyter. https://doi.org/10.1515/9783110719192-018

Joho, Tobias. 2022. *Style and Necessity in Thucydides.* Oxford, New York: Oxford University Press. https://doi.org/10.1093/oso/9780198812043.001.0001

Jouanna, Jacques. 1984. Rhétorique et Médecine dans la collection Hippocratique. Contribution à l'histoire de la rhétorique au Ve siècle. *Revue des Études Grecques* 97. 26–44.

— 2012. *Greek Medicine from Hippocrates to Galen: Selected Papers by Jacques Jouanna*. (Ed.) Neil Allies and Philip van der Eijk. Leiden; Boston, MA: Brill. https://doi.org/10.1163/9789004232549

Kalverkämper, Hartwig. 2005. Fachkommunikation zwischen Tradition und Innovation: Ein kulturhistorisches Phänomen der alten und modernen Gesellschaften. In *Antike Fachtexts/Ancient Technical Texts*, 319–362. Berlin; New York: De Gruyter. https://doi.org/10.1515/9783110912104.319

Karimullah, Kamran I. 2020. Hippocrates Transformed: Crafting a Hippocratic Discourse of Medical Semiotics in English, 1850–1930. *Humanities and Social Sciences Communications* 7(1). 27. https://doi.org/10.1057/s41599-020-0511-7

Kraska-Szlenk, Iwona. 2014. Semantic Extensions of Body Part Terms: Common Patterns and Their Interpretation. *Language Sciences* 44. 15–39. https://doi.org/10.1016/j.langsci.2014.02.002

Langholf, Volker. 1990. *Medical Theories in Hippocrates: Early Texts and the "Epidemics".*

Berlin; Boston, MA: De Gruyter.

Langslow, David. 2000. *Medical Latin in the Roman Empire.* Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198152798.001.0001

— 2012. The Language of Polybius since Foucault and Dubuisson. In Christopher Smith and Liv Mariah Yarrow (eds.), *Imperialism, Cultural Politics, and Polybius*, 85–110. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199600755.003.0006

Lateiner, Donald. 1986. The Empirical Element in the Methods of Early Greek Medical Writers and Herodotus: A Shared Epistemological Response. *Antichthon* 20. 1–20. https://doi.org/10.1017/s0066477400003415

Leitner, Magdalena, and Andreas H. Jucker. 2021. Historical Sociopragmatics. In Dániel Z. Kádár, Marina Terkourafi, and Michael Haugh (eds.), *The Cambridge Handbook of Sociopragmatics*, 687–709. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108954105.035

Lynch, Richard. 2014. Discourse. In John Nale and Leonard Lawlor (eds.), *The Cambridge Foucault Lexicon*, 120–125. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139022309

Lyons, John. 1995. *Linguistic Semantics: An Introduction*. Cambridge: Cambridge University Press.

Maclean, Ian. 2021. The Reception of Hippocrates by Physicians at the End of the Seventeenth Century. In Dmitri Levitin and Ian Maclean (eds.), *The Worlds of Knowledge and the Classical Tradition in the Early Modern Age: Comparative Approaches*, 217–274. Leiden; Boston, MA: Brill. https://doi.org/10.1163/9789004462335_009

Mader, Gottfried. 2018. Ἐναργὲς καὶ σαφές: Demosthenes and the Rhetoric of Disclosure in the Philippic Orations. *American Journal of Philology* 139(2). 177–214. https://doi.org/10.1353/ajp.2018.0011

Mambrini, Francesco. 2019. Nominal vs Copular Clauses in a Diachronic Corpus of Ancient Greek Historians. *Journal of Greek Linguistics* 19(1). 90–113. https://doi.org/10.1163/15699846-01901003

Mandēlaras, Vasileios. 1973. *The Verb in the Greek Non-literary Papyri.* Athens: Hellenic Ministry of Culture and Sciences.

Marušić, Borislav. 2018. Besonderheiten der Funktionsverbgefüge in der deutschen Konzernsprache. *Jezikoslovlje* 19(1). 87–106.

Mcarthur, Tom, Jacqueline Lam-Mcarthur, and Lise Fontaine. 2018. *Oxford English Dictionary*. Oxford: Oxford University Press.

Mel'čuk, Igor. 1996. Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon. In Leo Wanner (ed.), *Lexical Functions in Lexicography and Natural Language Processing*, 37–102. Amsterdam: John Benjamins.

— 2004. Verbes supports sans peine. *Lingvisticæ Investigationes* 27(2). 203–217. https://doi.org/10.1075/li.27.2.05mel

Miller, D. Gary. 2013. *Ancient Greek Dialects and Early Authors: Introduction to the Dialect Mixture in Homer, with Notes on Lyric and Herodotus.* Boston, MA: Mouton De Gruyter.

Montalt, Vicent, and Isabel García-Izquierdo. 2025. Bridging the Gaps between Medicine, Health Care and Society: Travelling Conceptualizations of Translation and Mediation. *Translation in Society* 4(1). 77–96. https://doi.org/10.1075/tris.24009.mon

Morison, Ben. 2008. Language. In Robert Hankinson (ed.), *The Cambridge Companion to Galen*, 116–156. Cambridge: Cambridge University Press. https://doi.org/10.1017/ccol9780521819541.005

Nelson, Eric. 2016. Tracking the Hippocratic Woozle: Pseudepigrapha and the Formation of the Corpus. In Lesley Dean-Jones and Ralph Rosen (eds.), *Ancient Concepts of the Hippocratic: Papers Presented at the XIIIth International Hippocrates Colloquium, Austin, Texas, August 2008*, 117–140. Leiden; Boston, MA: Brill. https://doi.org/10.1163/9789004307407_008

Nicholson, Emma. 2021. Polybios and the Rise of Rome: Gramscian Hegemony, Intellectuals, and Passive Revolution. In Emilio Zucchetti and Anna Maria Cimino (eds.), *Antonio Gramsci and the Ancient World*, 141–164. London: Routledge. https://doi.org/10.4324/9780429201684

Nicolai, Roberto. 2007. The Place of History in the Ancient World. In John Marincola (ed.), *A Companion to Greek and Roman Historiography*, 13–26. Malden: Blackwell. https://doi.org/10.1002/9781405185110.ch1

Paavilainen, Helena. 2009. *Medieval Pharmacotherapy Continuity and Change: Case Studies from Ibn Sīnā and Some of his Late Medieval Commentators*. Leiden; Boston, MA: Brill. https://doi.org/10.1163/ej.9789004171190.i-796

Panagl, Oswald. 1982. Produktivität in der Wortbildung von Corpussprachen: Möglichkeiten und Grenzen der Heuristik. *Folia Linguistica* 16(1–4). 225–240. https://doi.org/10.1515/flin.1982.16.1-4.225

Pierrehumbert, Janet. 2006. The Next Toolkit. *Journal of Phonetics* 34. 516–530. https://doi.org/10.1016/j.wocn.2006.06.003

Potter, Paul. 1988. *Hippocrates. Affections. Diseases 1. Diseases 2*. Cambridge, MA: Harvard University Press.

Probert, Philomen. 2019. Greek Dialects in the Lexicon. In Christopher Stray, Michael Clarke, and Joshua T. Katz (eds.), *Liddell and Scott: The History, Methodology, and Languages of the World's Leading Lexicon of Ancient Greek*, 200–225. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198810803.003.0012

Reggiani, Nicola. 2019. *Greek Medical Papyri: Text, Context, Hypertext.* Berlin; Boston, MA: Mouton De Gruyter. https://doi.org/10.1515/9783110536409

Rochette, Bruno. 2005. Die Übersetzung von Fachbegriffen bei Apuleius. In *Antike Fachtexts/Ancient Technical Texts*, 289–318. Berlin; New York: De Gruyter. https://doi.org/10.1515/9783110912104.289

Rood, Timothy. 2012. Polybius. In Irene de Jong (ed.), *Space in Ancient Greek Literature*, 179–197. Berlin; Boston, MA: Brill. https://doi.org/10.1163/9789004224384_011

— 2014. Space and Landscape in Xenophon's Anabasis. In Kate Gilhuly and Nancy Worman (eds.), *Space, Place, and Landscape in Ancient Greek Literature and Culture*, 63–93. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107323766.003

Roselli, Amneris. 2016. The Gynaecological and Nosological Treatises of the Corpus Hippocraticum: The Tip of an Iceberg. In Giulio Colesanti and Laura Lulli (eds.), *Submerged Literature in Ancient Greek Culture: Volume 2 Case Studies*, 187–204. Berlin: De Gruyter. https://doi.org/10.1515/9783110428636-013

Rusten, Jeffrey. 2020. τὴν ἐκβολὴν τοῦ λόγου ἐποιησάμην: Thucydides' Chronicle in

the Pentekontaetia (1.97–117) Is Not a Digression. *Histos* 14. 230–254. https://doi.org/10.29173/histos446

Sabanés, Laia Vidal, and Iria da Cunha. 2025. AI as a Resource for the Clarification of Medical Terminology: An Analysis of its Advantages and Limitations. *Terminology, International Journal of Theoretical and Applied Issues in Specialized Communication* 31(1). 37–71. https://doi.org/10.1075/term.00083.vid

Santis, Annamaria de, and Irene Rossi. 2019. *Crossing Experiences in Digital Epigraphy: From Practice to Discipline*. Berlin; Boston, MA: Mouton De Gruyter. https://doi.org/10.1515/9783110607208

Schepens, Guido. 2007. History and Historia: Inquiry in the Greek Historians. In John Marincola (ed.), *A Companion to Greek and Roman Historiography*, 39–55. Malden: Blackwell. https://doi.org/10.1002/9781405185110.ch3

Sheinfux, Livnat, Tali Greshler, Nurit Melnik, and Shuly Winter. 2019. Verbal Multiword Expressions: Idiomaticity and Flexibility. In Yannick Parmentier and Jakub Waszczuk (eds.), *Representation and Parsing of Multiword Expressions*, 35–68. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.2579035

Squeri, Elena. 2024. χράομαι as a Support Verb in the Medical Jargon of the Hippocratic Corpus. In Victoria Fendel (ed.), *Support-verb Constructions in the Corpora of Greek: Between Lexicon and Syntax?*, 133–164. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.14017929

Tagliamonte, Sali A. 2025. *Analysing Sociolinguistic Variation.* 2nd edn. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009403092

Thomas, Rosalind. 2000. *Herodotus in Context: Ethnography, Science and the Art of Persuasion.* Cambridge: Cambridge University Press. https://doi.org/10.2307/1192607

Totelin, Laurence. 2012. And to End on a Poetic Note: Galen's Authorial Strategies in the Pharmacological Books. *Studies in History and Philosophy of Science* Part A 43(2). 307–315. https://doi.org/10.1016/j.shpsa.2011.12.019

Trenery, Claire. 2020. Mind-Body Interaction: The Influence of Ancient Ideas in Twelfth-Century England. In Chiara Thumiger (ed.), *Holism in Ancient Medicine and Its Reception*, 377–399. Boston, MA; Leiden: Brill. https://doi.org/10.1163/9789004443143_017

Tronci, Liana. 2017. At the Lexicon-syntax Interface Ancient Greek Constructions with ἔχειν and Psychological Nouns. In Thanasis Georgakopoulos, Theodossia-Soula Pavlidou, Miltos Pechlivanos, Artemis Alexiadou, Jannis Androutsopoulos, Alexis Kalokairinos, Stavros Skopeteas, and Katerina Stathi (eds.), *Proceedings of the ICGL12*, Vol. 2, 1021–1033. Berlin: Edition Romiosini/CeMoG.

Tzanoulinou, Diamanto, Loukas Triantafyllopoulos, and Vassilios S. Verykios. 2025. Harnessing Language Models for Studying the Ancient Greek Language: A Systematic Review. *Machine Learning and Knowledge Extraction* 7(3). https://doi.org/10.3390/make7030071

Unwin, Lorna, Jason Hughes, and Nick Jewson. 2007. *Communities of Practice: Critical Perspectives.* London: Routledge. https://doi.org/10.4324/NOE0415364737

Vieregge, Annika. 2025. *Bewertung und Variation der Präpositionalkasus im Deutschen.* Berlin: Language Science Press. https://doi.org/10.5281/zenodo.14608443

Viti, Carlotta. 2024. Introduction: Language Contact, Comparative Linguistics, and Comparative Literature in their Historical and Cultural Context. In Carlotta Viti (ed.),

Ancient Greek and Latin in the Linguistic Context of the Ancient Mediterranean, 11–60. Tübingen: Narr.

Vives Cuesta, Alfonso, and Lucía Madrigal Acero. 2022. Support-verb Constructions in Postclassical Greek and Sociolinguistics: A Diachronic Study of εὐχὴν ποιέω as a Level-of-speech Marker. In María Dolores Jiménez López, María Jiménez Martínez, Cristina Tur Altarriba, and José Miguel Baños (eds.), *Collocations in Theoretical and Applied Linguistics: From Classical Languages to Romance Languages*, 305–334. Madrid: SEEC.

Wahlgren, Staffan. 2024. Word Order in Byzantine Literary Greek. In Giuseppina di Bartolo and Daniel Kölligan (eds.), *Postclassical Greek: Problems and Perspectives*, 293–308. Berlin; Boston, MA: Mouton De Gruyter. https://doi.org/10.1515/9783111079172-012

Wenger, Etienne. 1999. *Communities of Practice: Learning, Meaning, and Identity.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511803932

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3(1). https://doi.org/10.1038/sdata.2016.18

Willi, Andreas. 2003. *The Languages of Aristophanes: Aspects of Linguistic Variation in Classical Attic Greek*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199215102.001.0001

Wittenberg, Eva, and Roger Levy. 2017. If You Want a Quick kiss, Make It Count: How Choice of Syntactic Construction Affects Event Construal. *Journal of Memory and Language* 94. 254–271. https://doi.org/10.1016/j.jml.2016.12.001

Wright, Alex. 2014. *Cataloging the World: Paul Otlet and the Birth of the Information Age*. New York: Oxford University Press.

4. Old Egyptian multi-word expressions in Unas' and Teti's Pyramid Texts (24th to 23rd c. BCE)

Roberto A. Díaz Hernández

1. Introduction

The Pyramid Texts are a collection of spells recorded on the walls of the pyramids of the following Old Kingdom kings and queens: Unas, Teti, Pepi I, Merenre, Pepi II, Ankhenespepi II, Neith and Ibi. The earliest evidence of these texts is found in the pyramid of Unas (ca. 2321–2306 BCE). The Pyramid Texts were recited during the mortuary rituals carried out for the deceased and they are written in Old Egyptian (ca. 2700–2000 BCE), the earliest stage of the Egyptian language for which extensive documentation exists. They have been edited by Sethe (1908–1922) and more recently by Allen (2013).

Since Sethe's edition, the Pyramid Texts have been a valuable resource for studying Old Egyptian, with Edel (1955–1964) using many examples from the Pyramid Texts in his Old Egyptian Grammar, and Allen (1984) carrying out thorough analyses of the verb forms used in the Pyramid Texts. Allen (2017) himself has recently begun publishing a grammar for each corpus of Pyramid Texts.[1] In addition, Hainline (2026) has conducted a study of linguistic variation in the Pyramid Texts. A morphosyntactic annotation of the Pyramid Texts has also been underway in the Universal Dependencies Egyptian-Pre-Coptic treebank (EPC treebank henceforth) since March 2024.[2] This treebank currently contains 34,234 tokens and 3,089 sentences from the Pyramid Texts,[3] of which 11,664 tokens are from Unas' Pyramid Texts and 8,648 are from Teti's Pyramid Texts. This chapter provides a morphosyntactic analysis of the multi-word

1 The first volume is devoted to the Old Egyptian grammar in Unas's Pyramid Texts.
2 This treebank was originally named the "Egyptian-UJaen Treebank" (see Díaz Hernández and Passarotti 2024).
3 UD release 2.18, May 2026 (https://universaldependencies.org/).

©2026 Roberto A. Díaz Hernández, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.04

expressions (MWEs henceforth) that appear in Unas' and Teti's Pyramid Texts. As part of a project devoted to the synchronic and diachronic use of Egyptian MWEs, the MWEs in the Pyramid Texts of other kings are planned to be studied (Díaz Hernández 2024a: 31; 2024b: 112).

An MWE is a (continuous or discontinuous) sequence of words which should contain at least two lexicalised component words and display some degree of lexical, morphological, syntactic and/or semantic idiosyncrasy (Savary et al. 2018: 92), for example "to go bananas", meaning "to become angry".[4] Egyptian MWEs were identified and categorised according to the PARSEME 2.0 guidelines,[5] which distinguish four main types of MWE acting as specific parts of speech:

1. A verbal multi-word expression (VMWE henceforth), which consists of a verb as its head, for example *to **spill** the beans*.
2. A nominal multi-word expression (NMWE henceforth), which functions as a noun, for example ***hot dog***.
3. An adjectival or adverbial multi-word expression (AMWE henceforth), which functions as an adjective or adverb respectively, for example ***ice-cold** drinks* and ***from time to time***.
4. A functional multi-word expression (FMWE henceforth), which syntactically acts like a function word—i.e. a determiner, adposition, conjunction, or interjection—for example ***in front of***.

Each of these categories comprises subcategories of MWEs which are presented at the beginning of each section below. The subcategories found in Unas' and Teti's Pyramid Texts are explained using examples taken from the EPC treebank. This treebank was recently used to annotate Egyptian MWEs during the PARSEME 2.0 annotation campaign,[6] which took place from March to April 2025. A total of 633 MWEs were identified in the EPC treebank during this campaign (Savary et al. forthcoming: Table 2). Of these, 338 were found in Unas' Pyramid Texts and 235 in Teti's Pyramid Texts. This chapter analyses the use of MWEs in these two textual corpora as a case study, based on the results achieved during the PARSEME 2.0 annotation campaign.

2. MWEs in Unas' and Teti's Pyramid Texts

Table 4.1 shows the distribution of MWEs in Unas' and Teti's Pyramid Texts. MWEs appear more frequently in Unas' Pyramid Texts than in Teti's, except for inherently reflexive verbs, inherently adpositional verbs, and conjunction

4 This definition differs slightly from that of Baldwin and Kim's (2010: 274–279).

5 See https://parsemefr.lis-lab.fr/parseme-st-guidelines/2.0/

6 PARSEME (PARSing and Multi-word Expressions) is an international community devoted to develop and maintain a multilingual corpus for MWEs. It originated from a homonymous PARSEME COST Action (2013–2017) and has evolved within the UniDive COST Action (2022–2026). See https://gitlab.com/parseme/corpora/-/wikis/home

idioms, which appear slightly more frequently in Teti's. It should be noted that NMWEs are far more prevalent in Unas' Pyramid Texts than in Teti's. MWEs consisting of nisba adjectives[7] (see Sections 4.4 and 5.4) are included in nominal idioms and adjectival idioms. Idiomatic verb-particle constructions are not found in Egyptian. Similarly, there are no examples of multi-verb constructions in the two corpora, even though this MWE type occurs in other Egyptian texts.

MWE type	**Unas**	**Teti**
VMWE		
Verbal idiom (VID)	83	65
Light-verb construction (LVC.full)	9	7
Causative LVC (LVC.cause)	6	5
Inherently reflexive verb (IRV)	3	4
Idiomatic verb-particle construction (IVPC)	0	0
Multi-verb construction (MVC)	0	0
Inherently adpositional verb (IAV)	6	7
NMWE		
Nominal idiom (NID)	42	18
Pronominal idiom (PronID)	26	4
Deverbal nominal (NV)	24	7
AMWE		
Adjectival idiom (AdjID)	10	8
Adverbial idiom (AdvID)	30	25
Deverbal AMWE	9	2
FMWE		
Determiner idiom (DetID)	2	0
Adposition idiom (AdpID)	75	61
Conjunction idiom (ConjID)	3	7
Interjection idiom (IntjID)	10	15
Total	338	235

Table 4.1 Distribution of MWEs in Unas's and Teti's Pyramid Texts.

7 Nisba adjectives are formed by adding an ending to nouns, and rarely to prepositions and pronouns in Semitic languages, such as Arabic.

3. Verbal multi-word expressions

According to PARSEME guidelines 2.0, there are six types of VMWEs:

1. Verbal idioms (VIDs), for example *to go bananas*;
2. Light-verb constructions (LVCs), characterised by the use of a verb with a predicative noun. If the verb contributes only morphological features to the meaning of the whole, it is a full LVC, for example *to pay a visit*. If the verb has a causative meaning, it is a causative LVC, for example *to give a headache*;
3. Inherently reflexive verbs (IRVs). These can be classified into two types: those that are always used with a reflexive pronoun, such as the Spanish verb *suicidarse* (“to commit suicide”), and those that are occasionally used with a reflexive pronoun to express a different meaning to that of the original verb without a reflexive pronoun, for example *to help oneself* means to take something without asking, which is different from the meaning of the verb *to help*;
4. Idiomatic verb-particle constructions (IVPCs), which can be full if the particle changes the meaning of the verb, for example *to do in*, or semi non-compositional if the particle adds a predictable meaning to the verb, for example *to eat up*;
5. Multi-verb constructions (MVCs), for example *to let go*;
6. Inherently adpositional verbs (IAVs), which are characterised by the use of a preposition that is either always required or, if absent, changes the meaning of the verb, for example *to come across*.

3.1 Verbal idioms

The use of VIDs is common in Unas’ and Teti’s Pyramid Texts, often consisting of a verb and a body-part noun,[8] which may be: *ỉb* (“heart”) (18x); *ꜥw* (“arm”) (7x); *rʾ* (“mouth”) (2x); *rṭ* (“foot, leg”) (2x); *ḥɜ.t(i)* (“heart”) (3x); *ḥr* (“face”) (5x); *śɜ* (“back”) (1x); and *ḏr.t* (“hand”) (1x), for example:

(1) Pyramid Texts 741e Teti (PT 741e T)



sb	*=č*	*ꜥw*	*=f*	*n*	*Rꜥw*	*r*	*ɜḫ.t*
send:SBJV	=2SG.F	arm-M.SG	=3SG.M	to:PREP	Re-GN	on:PREP	horizon-F.SG

8 For the use of body-part MWEs in Old Egyptian, see Díaz Hernández (2024a).

LT:[9] "You shall send his hand for Re to the horizon."

FT: "You shall guide him to Re at the horizon."

Comment: The verb *sbi̯* usually means "to go" or "to send". According to the *Wörterbuch* (Wb henceforth), its meaning changes when used with the noun *ꜥw* "arm" to form the expression *sbi̯ ꜥw*, meaning "to guide somebody" (Wb. III 431, 27). This is indeed a VMWE, as replacing *sbi̯* or *ꜥw* with any other related words would result in an ungrammatical expression or an unexpected meaning, for example *sbi̯ imꜣḫ*, meaning "to reach the dignity" (German "die Zeit der Ehrwürdigkeit durchleben", see Wb. III 431, 24).

Unas' and Teti's Pyramid Texts also contain VIDs consisting of words other than body-part nouns used as modifiers. These words are: *wꜣ.t* ("way, path") (4x); *mṯw* ("word") (4x); *nkn* ("injury") (1x); *ḫrw* ("voice") (2x); *sꜣ* ("protection") (1x); *ś.t* ("place") (1x); *śḏ.t* ("fire") (1x); and *tꜣ (*"earth, ground") (1x), for example:

(2) Pyramid Texts 347b Teti (PT 347b T)







wḏꜥ	*=f*	*mṯw*	*nčr*	*iś*
judge:FUT	=3SG.F	word-M.SG	god-M.SG	as:POSTP







śḏm.n	*=f*	*mṯw*	*śr*	*iś*
hear:PST	=3SG.F	word-M.SG	official-M.SG	as:POSTP

LT: "He will judge a word as a god, (after) he has heard a word as an official."

FT: "He will render judgment as a god, having heard the case as an official."

Comment: Two VMWEs occur in this instance: *wḏꜥ mṯw*, which literally means "to separate/judge a word", i.e. "to render judgment", and *śḏm mṯw*, which literally means "to hear a word", i.e. "to hear a case" or "to hold an inquiry". Although the modifier is *mṯw* ("word") in both cases, the meaning of the first VMWE changes completely due to the replacement of the verb *wḏꜥ* with *śḏm*.

3.2 Light-verb constructions

Full LVCs are rarely found in Unas' and Teti's Pyramid Texts. For example:

9 LT stands for literal translation and FT for free translation.

(3) Pyramid Texts 277a Unas (PT 277a W)



ỉr	*ś.t*	*n*	*Wniś*
make:IMP	place-M.SG	for-PREP	Unas-KN

LT: “Make a place for Unas.”

FT: “Take a seat for Unas.”

Comment: According to the *Wörterbuch*, *irị ś.t* means “to take a seat” (German “Platz nehmen”, see Wb. IV 6, 6). Here, the verb *ỉri* (“to make”) has little semantic content, in contrast to the predicative noun *ś.t* (“place, seat”) which expresses the action of sitting.

In contrast, causative LVCs are common in Unas’ and Teti’s Pyramid Texts. These can be classified according to the verb used as their head: those consisting of *irị̂* (“to do, to make”) (10x), *wṭị̂* (“to put”) (1x) and *rḏị̂* (“to give, to cause”) (4x). Compare:

(4) Pyramid Texts 500c Unas (PT 500c W)



ỉr	*=f*	*n*	*=k*	*iȝ*	*iȝ*
make:SBJV	=3SG.M	for-PREP	=2SG.M	adoration-M.SG	adoration-M.SG

LT: “He shall make for you adoration (on) adoration.”

FT: “He shall adore you over and over.”

(5) Pyramid Texts 302c Unas (PT 302c W)



ỉw	*(w)ṯ.n*	*Wniś*	*nr.w*	*=f*	*m*	*ỉb*	*=śn*
AUX	put:PST	Unas-KN	fear-M.SG	=3SG.M	in:PREP	heart-M.SG	=3PL

LT: “Unas has put his fear in the heart of them.”

FT: “Unas has instilled fear in them.”

(6) Pyramid Texts 197d Unas (PT 197d W)



ḏi	*=č*	*mrw.t*	*Wniś*	*pn*
give:SBJV	=2SG.F	love-M.SG	Unas-KN	DEM:M.SG

LT: "You should give the love of this Unas."

FT: "You should spread Unas' love."

In causative LVCs, there is a semantic difference between *irị* "to do, to make" and *rḏi* "to give"—the former emphasises the realisation of a fact, i.e. a factitive action, while the latter focusses on the cause of an action. This becomes clear when *irị* "to do, to make" and *rḏi* "to give" are used with the same noun. Compare:

(7) Pyramid Texts 737d Unas (PT 737d W)



ir	*=ś*	*ki.t*	*=k*	*ḫr*	*nčr(.w)*
make: sbjv	=3sg.f	acclaim-m.sg	=2sg.m	before:prep	gods-m.pl

LT: "It (the Eye of Horus) shall make your acclaim before the gods."

FT: "It (the Eye of Horus) shall acclaim you before the gods."

(8) Pyramid Texts 197c Unas (PT 197c W)



ḏi	*=č*	*ki.t*	*Wniś*	*pn*
give:sbjv	=2sg.f	acclaim-m.sg	Unas-kn	dem:m.sg

LT: "You should give the acclaim of this Unas."

FT: "You should spread Unas's acclaim."

3.3 Inherently reflexive verbs

As the verb *ś:ms* "to proceed" is only attested with the reflexive pronoun *św* "himself" (Pyramid Texts 10c Queen Neith (PT 10c N); see also Wb. V 141, 14), it can be considered as an IRV that is always formed with a clitic. Unas' and Teti's Pyramid Texts also contain the following IRVs which have a different meaning

to the original verb without a reflexive pronoun: *i:ꜥbꜣ čw* (“to make yourself ready”), cf. *ꜥbꜣ* (“to provide”); *ś:mn św* (“to establish himself”), cf. *ś:mn* (“to make firm”); *i:ms kw* (“to betake yourself”), Wb. II 135, cf. *ms* (“to bring”); and *čs.y =f św* (“he may raise himself, i.e. get up”), Wb. V 406, cf. *čsi̯* (“to lift up”). For example:

(9) Pyramid Texts 317c Unas (PT 317c W)



čs.y	*=f*	*św*
raise:SBJV	=3SG.M	3SG.M

LT: “He shall raise himself.”

FT: “He shall get up.”

(10) Pyramid Texts 294c Teti (PT 294c T)



čs.y	*Tti*	*pn*	*nn*
lift up:FUT	Teti-KN	DEM:M.SG	DEM:SG

“This Teti will lift up this.”

Comment: The literal meaning of *čsi̯* is “to lift up” something, as in (10). However, when used with a reflexive pronoun, it takes on the figurative meaning of “to lift oneself up” (i.e. “to get up”), as in (9).

3.4 Multi-verb constructions

There are no instances of multi-verb constructions in Unas’ and Teti’s Pyramid Texts. However, this type of MWE is used with verbs of motion in Old Egyptian, for example:

(11) Siege *et al.* (2024: 80–82)



šm<.t>	*i̯w.t*
go:INF	come:INF

LT: “Going and coming” (Fischer 1977: 115–119).

FT: “A visit.”

3.5 Inherently adpositional verbs

Unas' and Teti's Pyramid Texts also provide instances of inherently adpositional verbs, characterised by the use of a specific preposition which changes their original meaning, for example *irị r* (" to act against (sb.)"), cf. *irị* ("to do, to make"), *pḫr n* ("to serve (sb.)"), cf. *pḫr* ("to turn round"), and *śḏm n* ("to listen to (sb.), to obey"), cf. *śḏm* ("to hear"):

(12) Pyramid Texts 417a Unas (PT 417a W)





ir	*=k*	*ir*	*=k*
do:SBJV	=2SG.M	to:PREP	=2SG.M

LT: "You shall do to you."

FT: "You shall act against yourself."

(13) Pyramid Texts 135c Unas (PT 135c W)




pḫr	*n*	*=k*	*iꜣ.(w)t*	*Ḥr.w*
turn:FUT	to:PREP	=2SG.M	mound-F.PL	Horus-GN

LT: "The mounds of Horus will turn to you."

FT: "The mounds of Horus will serve you."

(14) Pyramid Texts 611b Teti (PT 611b T)





śḏm	*n*	*Ḥr.w*
hear:IMP	to:PREP	Horus-GN

LT: "Hear to Horus!"

FT: "Listen to Horus!"

4. Nominal multi-word expressions

The PARSEME guidelines 2.0 distinguish three types of NMWEs:

1. Nominal idioms (NIDs), e.g. *hot dog*;
2. Pronominal idioms (PronIDs), e.g. *we love each other*;
3. Deverbal nominal MWE (NV), e.g. *heart breaker*, derived from the VID *to break one's heart*.

Additionally, there are nominal idioms consisting of nisba adjectives used as nouns in Old Egyptian (see Section 4.4).

4.1 Nominal idioms

In Unas' and Teti's Pyramid Texts, body-part nouns are usually used as modifiers in nominal idioms, for example:

(15) Pyramid Texts 208b Unas (PT 208b W)







šn	*čw*	*ḫnw-*	*ꜥw.w(i)*	*mw.t*	*=k*
enfold:SBJV	2SG.M	interior-M.SG	arm-M.DU	mother-F.SG	=2SG.M

LT: "The interior of the arms of your mother shall enfold you."

FT: "The embrace of your mother shall enfold you."

Comment: The word *ꜥw.w(i)* is the dual of *ꜥw* ("arm") and it cannot change its number when used with *ḫnw* ("interior") in the MWE *ḫnw-ꜥw.w(i)* ("embrace") (Wb. III 372, 6).

Other body-part nouns are also used in NIDs, for example:

1.  *tꜣ-rṭ(.w)* ("ground (for) feet", i.e. "stairway") (PT 365a W);
2.  *ś.t-ib* ("place (of the) heart", i.e. "favourite place") (PT 414c W);
3. *sꜣ n(.i) ḫ.t* ("son of (my) belly", i.e. "(my) bodily son") (PT 1d T).

NIDs without body-part nouns also occur in Unas' and Teti's Pyramid Texts, for example:

1.  *mṭw-nčr* ("god's word", i.e. "sacred writings") (PT 333c T, see also Wb. II 181, 1);
2.  *bw-mꜣꜥ* ("the place of truth", i.e. "correctness, justice") (PT 4c T, see also Wb. II 14, 12);
3.  *ś.t ḥtp.w* ("place of peace", i.e. "place of rest") (PT 508b W, see also Wb. IV 5, 10).

4.2 Pronominal idioms

The following pronominal idioms occur in Unas' and Teti's Pyramid Texts: *pw-nn* ("this here",[10] i.e. "this one"); *wᶜ pn* ("one, this", i.e. "this one"); *wᶜ =śn(i)* ("one of both", i.e. "one of them"); *č.t =ś* ("her body", i.e. "herself"); and *čś =f* ("his person", i.e. "himself"):

(16) Pyramid Texts 167a Unas (PT 167a W)



sꜣ	*=k*	*pw*	*p(w)-*	*nn*
son-M.SG	=2SG.M	COP	DEM:M.SG	here:ADV

LT: "Your son is this here."

FT: "Your son is this one."

(17) Pyramid Texts 254a Unas (PT 254a W)



wᶜ	*pn*	*im(.i)*	*=k*
one-M.SG	DEM:M.SG	one who is in -NISBA.NOM	=2SG.M

LT: "This one, one who is in you."

FT: "This one who is you."

(18) Pyramid Texts 472d Unas (PT 472d W)



wᶜ	*=śn(i)*	*m*	*pn*	*gś*
one-M.SG	=3DU	on:PREP	DEM:M.SG	one-M.SG

LT: "One of both (is) on this side."

FT: "One of them is on this side."

10 The demonstrative pronoun *pw* forms a fixed connection with the adverb *nn*, see Edel (1955–1964: 491 (§ 963)). Since *pw* "this" is the head of this MWE, *pw-nn* can be considered as a pronominal idiom meaning "this one".

(19) Pyramid Texts 594c Teti (PT 594c T)









i:nḏ	*=ś*	*ḏ.t*	*=ś*	*m-*	*ꜥw*	*Śtḫ*
protect:SBJV	=3SG.F	body-F.SG	=3SG.F	from:PREP	arm-F.SG	Seth-GN

LT: “It (i.e. the Eye of Horus), its body, shall protect from the arm of Seth.”

FT: “It shall protect itself from Seth.”

(20) Pyramid Texts 529b Teti (PT 529b T)








Tti	*ḏś*	*=f*	*m-*	*ṯp*	*hrw*
Teti-KN	person-M.SG	=3SG.M	in:PREP	head-M.SG	day-M.SG

LT: “Teti, his person, (is) [in] the head of the day.”

FT: “Teti himself is at daybreak.”

4.3 Deverbal nominal MWEs

The head of deverbal nominal MWEs is usually a participle used as a noun, for example, *ḫśr-nčr* (“one who repels the god”, i.e. “scorpion”) (Pyramid Texts 227b Unas (PT 227b W)). However, it can also be an infinitive or a noun derived from a verb, for example, *pr(.t)-ḫrw* (“going out of the voice”, i.e. “invocation offering”) (Pyramid Texts 59d Unas (PT 59d W)) and *iꜥ.w-rʾ* (“washing the mouth”, i.e. “breakfast”) (Pyramid Texts 60a Unas (PT 60a W)). There are three types of deverbal nominal MWEs in Unas’ and Teti’s Pyramid Texts. Deverbal nominal derived from a VID (NV.VID), for example:

(21) Pyramid Texts 707c Teti (PT 707c T)



nḏm-	*ib*
be sweet:inf	heart-m.sg

LT: “Sweetness of heart”

FT: “Happiness.”

Deverbal nominal derived from an LVC (NV.LVC):

(22) Pyramid Texts 326d Teti (PT 326d T)







Wniś	*pi*	*ir*	*wp.t*	*nšn*
Unas-KN	COP	do-PTCP.ACT(M.SG)	mission-F.SG	storm-M.SG

LT: "Unas is he who does the mission of the storm."

FT: "Unas is he who carries out the mission of the storm."

Deverbal nominal derived from an IAV (NV.IAV):

(23) Pyramid Texts 713b Teti (PT 713b T)








m-	*ꜥw*	*ir:yt*	*ir*	*=f*	*ḏw*
from:prep	arm-m.sg	do-ptcp.pass(f.sg)	to:prep	=3sg.m	bad:adv

LT: "(...) from the arm of what could be evilly done to him."

FT: "(...) from what could be evilly done against him."

4.4 Nisba idioms in nominal function

An adjectival MWE consisting of a nisba adjective becomes a nominal idiom when the nisba adjective is used as a noun, for example:

(24) Pyramid Texts 211a Unas (PT 211a W)




nỉ	*ỉm(.i)-*	*rṭ*	*=k*
not-NEG	one who is in-M.SG	foot-M.SG	=2SG.M

LT: "There is no one who is in your foot."

FT: "You have no enemy."

Comment: The nisba adjective *ỉm(.i)* ("one who is in") derives from the preposition *m* ("in"). It forms a compound with the noun *rṭ* ("foot"). The idiosyncratic meaning of *ỉm(.i)-rṭ* ("foot") results from the Egyptian custom of decorating sandals with images of enemies to be trodden upon.

Nisba idioms in nominal function are common in Unas' and Teti's Pyramid Texts, especially those consisting of body-part nouns used as modifiers, for example:

(25) Pyramid Texts 404b Unas (PT 404b W)





ḥr.(i)w-	*ib*	*=śn*
those who are on-m.pl	heart-m.sg	=3pl

LT: “Those who are on their hearts.”

FT: “Their middle-sized ones.”

(26) Pyramid Texts 256b Unas (PT 256b W)

n(.i)-	*ś(w)*	*=k*
one who concerns-m.sg	3sg.m	=2sg.m

LT: “One concerning him (relating to) you.”

FT: “Your relative.”

5. Adjectival and adverbial MWEs

Three types of adjectival and adverbial MWEs can be distinguished in the 2.0 PARSEME guidelines:

1. Adjectival idioms (AdjIDs), e.g. *a bottom-up algorithm*;
2. Adverbial idioms (AdvIDs), e.g. *by and large*;
3. Deverbal adjectival and adverbial MWEs MWE (AV), e.g. *a run-down apartment*, derived from the IVPC *to run down*.

Additionally, there are adjectival idioms consisting of nisba adjectives in Old Egyptian. Deverbal adverbial MWEs are unattested in Unas’ and Teti’s Pyramid Texts.

5.1 Adjectival idioms

The following adjectival sentence is used as an adjectival idiom in Unas’ and Teti’s Pyramid Texts:

(27) Pyramid Texts 195c Unas (PT 195c W)





nfr	*w(i)*	*ḥr*	*=č*
good- M.SG	PTCL	face-M.SG	=2SG.F

LT: "How goodly is your face."

FT: "How nice it is to see you."

5.2 Adverbial idioms

These adverbial MWEs usually consist of a preposition and a noun, for example *m-sp* "at once", (28), and *n ḏ.t* "for ever", (29):

(28) Pyramid Texts 590c Teti (PT 590c T)



in.n	*=f*	*n*	*=k*	*nčr(.w)*	*nb(.w)*	*m-*	*sp*
bring:PST	=3SG.M	to:PREP	=2SG.M	god-M.PL	every-M.PL	in:PREP	time-M.SG

LT: "He brought all the gods to you (certainly) in time."

FT: "It is at once that he brought all the gods to you."

(29) Pyramid Texts 160c Unas (PT 160c W)



sꜣ	*=k*	*pw*	*n(.i)*	*ḏ.t*	*=k*	*n*	*ḏ.t*
son-M.SG	=2SG.M	DEM:M.SG	of-M.SG	body-F.SG	=2SG.M	for:PREP	eternity-F.SG

LT: "Your son of your body is this for the eternity."

FT: "He is your bodily son for ever."

Adverbial idioms can also be noun phrases used as adverbs, for example *rˁw nb* "every day":

(30) Pyramid Texts 123k Teti (PT 123k T)



śḏr	*Tti*	*wḏꜣ(.w)*	*rˁw*	*nb*
lie down:SBJV	Teti-KN	be healthy:OSSC.3SG.M	sun-M.SG	each-M.SG

LT: "Teti shall lie down being in good health each sun."

FT: "Teti shall lie down in good health every day."

5.3 Deverbal adjectival MWEs

According to their morphosyntactic use in Unas' and Teti's Pyramid Texts, two types of deverbal adjectival MWEs can be distinguished: those consisting of a participle as an adjectival modifier and those consisting of a participle or a

verbal adjective modified by a noun (i.e. a *bahuvrīhi* construction, also known as a "limitation" in Egyptology) (Peust 2008: 64–66).[11] In the first type, the verb is used as a participle to modify an antecedent, for example:

(31) Pyramid Texts 511c Unas (PT 511c W)






pśḏ.(w)t	*śḏm.t*	*mṯw*	*iti.w*
Ennead- F.PL	hear-PTCP.ACT(F.SG)	word-M.SG	monarch-M.SG

LT: "(...) The Enneads which hear the word of the monarch."

FT: "(...) The Enneads which hear the case of the monarch."

Comment: The AMWE *śḏm.t mṯw* derives from the VMWE *śḏm mṯw* (cf. example (2) above).

A deverbal adjectival MWE can also consist of a participle or a verbal adjective derived from a verb used in a VMWE, for example:

(32) Pyramid Texts 507b Unas (PT 507b W)









Wniś	*pi*	*Śbk (...)*	*rś*	*ḥr*	*čs*	*ḥ3.t*
Unas-KN	COP	Sobek-GN	be watchful-PTCP.ACT(M.SG)	face-M.SG	raised-PTCP.ACT(M.SG)	brow-F.SG

LT: "Unas is Sobek (...), watchful of face, raised of brow."

FT: "Unas is Sobek (...), a vigilant and bold one."

Comment: In this example, *rś ḥr* derives from the VMWE *rś ḥr* ("watchful of face", i.e. "vigilant") (Wb. II 450, 17) and *čs ḥ3.t* from the VMWE *čsị ḥ3.t* ("raised of brow", i.e. "courageous" or similar).[12]

5.4 Nisba-adjective idioms

In Old Egyptian, nisba adjectives can be formed from nouns and prepositions by adding the *i*-ending. For example, the nisba adjective *n.i* ("belonging to") derives from the preposition *n* ("for"). An adverbial idiom consisting of a preposition and a noun can therefore be transformed into an adjectival idiom, by using the nisba adjective of the preposition. This occasionally occurs in Unas' and Teti's

11 This construction consists of an adjective as a head and a noun as its modifier.

12 The figurative meaning of *čsị ḥ3.t* is uncertain. The *Wörterbuch* provides only the literal meaning (Wb. III 19, 8).

Pyramid Texts, for example:

(33) Pyramid Texts 442b Unas (PT 442b W)







ḥfꜣ	*pn (...)*	*ḫr.i*	*ḏbꜥ(.w)*	*Wniś*
snake-M.SG	DEM:M.SG	who is under-M.SG	finger-M.PL	Unas-KN

LT: "(...) The snake (...) which is under Unas's fingers."

FT: "(...) The snake (...) which is under Unas's supervision."

Comment: This nisba adjectival idiom *ḫr.i ḏbꜥ(.w)* derives from the FMWE *ḫr ḏbꜥ(.w)* ("under the fingers", i.e. "under supervision") (Wb. V 565, 3–4).

As is common in Unas' and Teti's Pyramid Texts, a compound preposition can also be used to form a nisba adjective, for example:

(34) Pyramid Texts 2a Teti (PT 2a T)







Nw.t	*ꜥꜣ.t*	*ḥr.t-*	*ib*	*ḥw.t*	*ḫr.t*
Nut-GN	great-F.SG	who is on-F.SG	heart-M.SG	mansion-F.SG	which is low-F.SG

LT: "The great Nut, who is on the heart of the Lower Mansion."

FT: "The great Nut, who is in the middle of (i.e. dwells in) the Lower Mansion."

Comment: The nisba adjective idiom *ḥr.t-ib* derives from the compound preposition *ḥr-ib* ("on the heart", i.e. "in the middle of") (Wb. V 565, 3–4).

(35) Pyramid Texts 644e Teti (PT 644e T)






Ḥr.w	*ḥr(.i)-*	*ṯp*	*rḫ.(w)t*	*=f*
Horus-GN	who is on-M.SG	head-M.SG	subject-F.PL	=3SG.M

LT: "Horus who is on the head of his subjects."

FT: "Horus who rules over his subjects."

Comment: The nisba adjective idiom *ḥr(.i)-ṯp* derives from the compound preposition *ḥr-ṯp* ("on the head/top", i.e. "to rule over") (Wb. V 565, 3–4).

When a nisba-adjective idiom is used as a noun, it becomes a nisba idiom in nominal function (see Section 4.4).

6. Functional MWEs

According to the PARSEME guidelines 2.0, there are four types of functional MWEs:

1. Determiner idioms (DetIDs), e.g. *every other day*;
2. Adposition idioms (AdpIDs), e.g. ***in front of*** *the station*;
3. Conjunction idioms (ConjIDs), e.g. *she was fortunate* ***in that*** *she had friends to help her*;
4. Interjection idioms (IntjIDs), e.g. *damn it!*.

These types are all found in Unas' and Teti's Pyramid Texts.

6.1 Determiner idioms

Two determiner idioms are used in Unas' Pyramid Texts, whereas this MWE category is not found in Teti's Pyramid Texts:

(36) Pyramid Texts 460b Unas (PT 460b W)



nn	*n(.i)*	*rm.wt*	*nčr*
dem	of-m.sg	mourning-f.sg	god-m.sg

LT: "This of mourning of the god."

FT: "This mourning of the god."

(37) Pyramid Texts 264b Unas (PT 264b W)



nw	*n(.i)*	*sšsš*
dem	of-m.sg	lotus-m.sg

LT: "This of lotus."

FT: "This lotus."

Comment: The determiner idioms *nn n*(.*i*) and *nw n*(.*i*) show an idiosyncratic syntactic structure, consisting of the demonstratives *nn* and *nw* followed by the nisba adjective *n*(.*i*) ("belonging to", i.e. "of") (Edel 1955–1964: 89 (§ 199)).

6.2 Adposition idioms

In Old Egyptian, adposition idioms are exclusively compound prepositions, i.e. fixed expressions which usually consist of a preposition followed by a noun, for example:

(38) Pyramid Texts 735c Teti (PT 735c T)



(...) m-	*ꜥb*	*nčr(.w)*
in:PREP	association-M.SG	god-M.PL

LT: "(...) In association with the gods."

FT: "(...) In company with the gods."

(39) Pyramid Texts 477c Unas (PT 477c W)



(...) mr	*ḳṯ*	*iṯ.y*	*n*	*=f*	*im*
according to:PREP	form-M.SG	do-PTCP.PASS(M.SG)	for:PREP	=3SG.M	there:ADV

LT: "(...) According to the form which was done for him there."

FT: "(...) In accordance with how it was done for you."

The noun of a compound preposition is usually a body-part noun. There are twelve compound prepositions consisting of a body-part noun in Unas' and Teti's Pyramid Texts, for example:

(40) Pyramid Texts 270d Unas (PT 270d W)



(...) m-	*ṯp*	*kꜣ*
at:prep	head-m.sg	bull-m.pl

LT: "(...) At the head of the bull."

FT: "(...) Before the bull."

A compound preposition can occasionally be formed by the combination of a preposition, a nisba adjective derived from a preposition and a noun, for example:

(41) Pyramid Texts 582a Teti (PT 582a T)





(...) m-	*ḥr(.i)-*	*ib*	*=f*
in:PREP	one who is on -NISBA.NOM	heart-M.SG	=3SG.M

LT: “(...) In the one who is on his heart.”

FT: “(...) In his midst.”

Compound prepositions consisting of two prepositions are also used in the Pyramid Texts of Unas and Teti, for example:

(42) Pyramid Texts 130c Unas (PT 130c W)







ꜥ3b.t	*=f*	*m-*	*m*	*=čn*
food-offerings-F.SG	=3SG.M	in:PREP	in:PREP	=2PL

LT: “His food offerings are in-in you.”

FT: “His food offerings are among you.”

6.3 Conjunction idioms

Two conjunction idioms with a causal meaning are attested in Unas’ and Teti’s Pyramid Texts: *n-n.tt* (three times in Unas and once in Teti), meaning “for that which is” i.e. “on account of the fact that” (i.e. “because”); and *n-iw.t* (six times in Teti), meaning “for which not” (i.e. “because not”). The former consists of the preposition *n* (“for”) and the substantivised form of the relative pronoun *n.ti*, meaning “that which is” (Borghouts 2010: 123 (23.b.3); see also Schenkel 2012: 337–338). For example:

(43) Pyramid Texts 121b Unas (PT 121b W)







(...) n-	*n.tt*	*św.t*	*iś*	*kꜣ*	*wr*
for:PREP	that-which-is	3SG.M	PTCL	bull-M.SG	great-M.SG

LT: “(...) On account of the fact that he is certainly the Great Bull.”

FT: “(...) Because he is certainly the Great Bull.”

(44) Pyramid Texts 659d Teti (PT 659d T)



(...) n-	*iw.t*	*mw.t*	*=k*	*mś.ti*	*čw*	*m*	*rmč(.w)*
for:PREP	that-not	mother-F.SG	=2SG.M	bear: PRF-3SG.F	2SG.M	in:PREP	man-M.PL

LT: "(...) For which no mother of yours bore you among men."

FT: "(...) Because no mother of yours bore you among men."

Comment: The conjunction idiom *n-iw.t* is formed by combining the preposition *n* ("for") and the negative particle *iw.t* ("that not") derived from the adjective *iw.ti* ("who is without") (Borghouts 2010: 124 (23.b.3); see also Schenkel 2012, 337 and 339).

6.4 Interjection idioms

There are two common interjection idioms in Unas' and Teti's Pyramid Texts: *m=k* ("behold you!", i.e. "look!") and *(i)r=f* ("as to him", i.e. "indeed"):

(45) Pyramid Texts 333a Unas (PT 333a W)



m	*=k*	*Wniś*	*pr(.w)*
behold:imp	=2sg.m	Unas-kn	go:ossc.3sg.m

LT: "Behold you! Unas has gone up."

FT: "Look! Unas has gone up."

(46) Pyramid Texts 159b Unas (PT 159b W)



ii	*r*	*=f*	*Wniś*	*pn*
come:pst	as for:prep	=3sg.m	Unas-kn	dem:m.sg

LT: "As for him, this Unas came."

FT: "Indeed, this Unas came."

Comment: The interjection idiom m=k consists of m, probably an obsolete imperative meaning "behold!" (Gardiner 1957: 178 (§ 234)), and a second person suffix pronoun "you". It is used to draw the listener's attention. As for (*i*)r *=f*, this is originally formed by the preposition (*i*)r ("as to") and a suffix pronoun which varies depending on the context. This interjection idiom has an emphatic meaning (Gardiner 1957: 186 (§ 252)).

7. Conclusion

This study has revealed that most PARSEME categories of MWEs are used in Unas' and Teti's Pyramid Texts. Multi-verb constructions and idiomatic verb-particle constructions are the only exceptions. In contrast, Old Egyptian had two language-specific MWE types characterised by the use of nisba adjectives (see Sections 4.4. and 5.4).

Although MWEs are used uniformly in Unas' and Teti's Pyramid Texts, some differences have been found between the two textual corpora. For example, NMWEs are more common in Unas' Pyramid Texsots than in Teti's, and determiner idioms are only attested in the former. These variations may be attributed to compositional or stylistic factors rather than sociolinguistic issues, such as the use of dialectal expressions.

Finally, this work has paved the way for identifying Old Egyptian MWEs, enabling us to find and categorise more of them in the rest of the Pyramid Texts and other Egyptian texts once the morphosyntactic annotation of these texts in the EPC treebank is complete. This will contribute to the development of synchronic and diachronic studies on the use of Egyptian MWEs.

Acknowledgements

This work received support from the CA21167 COST action UniDive, funded by COST (European Cooperation in Science and Technology).

Bibliography

Allen, James P. 1984. *The Inflection of the Verb in the Pyramid Texts (2 vols.)*. Malibu, CA: Undena Publications.

— 2013. *A New Concordance of the Pyramid Texts (6 vols.)*. Providence, RI: Brown University.

— 2017. *A Grammar of the Ancient Egyptian Pyramid Texts. Volume I: Unis*. Winona Lake, IN: Eisenbrauns. https://doi.org/10.5325/j.ctv1bxh5bv

Baldwin, Timothy, and Su Nam Kim. 2010. Multiword Expressions. In Nitin Indurkhya and Fred J. Damerau (eds.), *Handbook of Natural Language Processing*, 267–292. Boca Raton, FL: Taylor and Francis. https://doi.org/10.1201/9781420085938-21

Borghouts, Joris F. 2010. *Egyptian. An Introduction to the Writing and Language of the Middle Kingdom (2 vols.)*. Leuven: Peeters.

Díaz Hernández, Roberto A. 2024a. Identification and Annotation of Body Part Multiword Expressions in Old Egyptian. In Archna Bhatia et al. (eds.), *Proceedings of the Joint Workshop on Multiword Expressions and Universal Dependencies (MWE-UD) @ LREC-COLING 2024*, 24–32. Turin: ELRA and ICCL.

— 2024b. Old Egyptian Multiword Expressions Consisting of a Head Word and ib "Heart". *Lingua Aegyptia* 32. 71–118. https://doi.org/10.37011/lingaeg.32.02

Díaz Hernández, Roberto A., Passarotti, Marco. C. 2024. Developing the Egyptian-UJaen treebank. In Daniel Dakota et al. (eds.), *Proceedings of the 22nd Workshop on Treebanks and Linguistic Theories (TLT 2024)*, 1–10. Hamburg: Association for Computational Linguistics.

Edel, Elmar. 1955–1964. *Altägyptische Grammatik*. Rome: Pontificium Institutum Biblicum.

Erman, Adolf, and Hermann Grapow (eds.). 1926–1961. *Wörterbuch der Aegyptischen Sprache*. Berlin: Hinrichs.

Fischer, Henry George. 1977. *The Orientation of Hieroglyphs*. New York: The Metropolitan Museum of Art.

Gardiner, Alan. H. 1957. *Egyptian Grammar. Being an Introduction to the Study of Hieroglyphs*. Oxford: Oxford University Press.

Hainline, Brendan. 2026. *"Speaking the Words": Linguistic Variation in the Pyramid Texts.* Leiden: Brill. https://doi.org/10.1163/9789004734876

Peust, Carsten. 2008. Adjektiv und Adjektivverb im Mittelägyptischen. In Carsten Peust (ed.), *Miscellanea in honorem Wolfhart Westendorf*, 58–82. Göttingen: Seminar für Ägyptologie und Koptologie der Universität Göttingen.

Savary, Agata et al. 2018. PARSEME Multilingual Corpus of Verbal Multiword Expressions. In Stella Markantonatou et al. (eds.), *Multiword Expressions at Length and in Depth. Extended Papers from the MWE 2017 Workshop*, 84–148. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.1471591

Savary, Agata et al. Forthcoming. PARSEME 2.0 Multilingual Corpus of Multiword Expressions.

Schenkel, Wolfgang. 2012. *Tübinger Einführung in die klassisch-ägyptische Sprache und Schrift*. N.p.: Tübingen.

Sethe, Kurt. 1908–1922. *Die altägyptischen Pyramidentexte nach den Papierabdrücken und Photographien des Berliner museums (4 vols.)*. Leipzig: Hinrichs.

II. METHODOLOGICAL NOVELTIES: CATEGORIES, CONTINUA, AND THEIR EXPANSION

5. Assessing verb-particle constructions through classical Greek preverbation (8th to 4th c. BCE)

Elena Squeri

1. Background, methodology, and purpose

This chapter tests the possibility of treating classical Greek pre-verbed verbs, for example, ἀπο-βάλλω *apo-ballō* "to throw away", and verb-particle constructions (VPCs henceforth), for example, English *to give away*, as instantiations of a cross-linguistic comparative concept (Haspelmath 2010).[1] It further aims to identify some first categories for a more thorough corpus study of classical Greek pre-verbed verbs while also employing the semantic and morphosyntactic data thus acquired to better define English VPCs.

1.1 Verb-particle constructions, phrasal verbs, and preverbation

The descriptive category of phrasal verb is peculiar to the English language (Dehé 2002: ch. 1; Thim 2012: ch. 2; Luo 2019: ch. 1). It refers to verb-particle combinations achieving a shared predication, with different degrees of idiosyncrasy in meaning, for example, *to go out* and *to give up*. The category of verb-particle constructions, considered sometimes synonymous (Dehé 2002: 1; McIntyre 2015) sometimes more inclusive (Luo 2019: §1.2) than the one of phrasal verbs, is well represented in Germanic languages (Thim 2012: §2.5), but is overall well attested typologically, also outside of the Indo-European family (Thim 2012: §3.1.1). The positioning of the particle before or after the verb is

1 Dataset: https://doi.org/10.5281/zenodo.18340160

©2026 Elena Squeri, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.05

not a separating feature, being influenced by the word order preferred by each (stage of a) language (Thim 2012: ch. 3).

The category of preverbs on the other hand is a "descriptive notion in Indo-European linguistics. It refers to morphemes that appear before a verb, and which form a close semantic unit with that verb. In many cases, the morpheme that functions as a preverb can also function without a preverbal context, often as an adverb or an adposition" (Booij and van Kemenade 2003: 1). Classical Greek preverbs occur before a verb but can also head prepositional phrases, their adverbial use being only attested in Homeric Greek.

Therefore, pre-verbed verbs instantiate in a single lexeme, while VPCs are realised as verb-governed multi-word expressions (MWEs henceforth), i.e. complex predicates (McIntyre 2015: §3). However, in older Indo-European languages, preverbs and verbs could still occur separately, their cohesion increasing diachronically (Pinault 1995). This stage is testified by Homeric Greek "tmesis" (McConnel 2025: ch. 5), often suggesting the parallel with "phrasal verbs" (Horrocks 1981; Pompei 2014).

(1) Homer, *Iliad*, 1.437

ἐκ	δὲ	καὶ	αὐτοὶ	**βαῖνον**
ek	*de*	*kai*	*autoi*	*bainon*
outside	prt	and	them.nom.pl.m	step.aor.ind.act.3pl

"And they all **stepped outside**."

(2) Homer, *Odyssey*, 7.284–285

ποταμοῖο	**ἐκβὰς**
potamoio	*ekbas*
river.gen.m.pl	step-out.aor.prtc.act.nom.sg.m

"**Stepping out** from the river."

In Classical Greek, verb and preverb are univerbated, but can still be separated by the augment (pres. ἐκβαίνω *ekbainō*—impf. ἐξέβαινον *ex-e-bainon*), which was by then a standardised tense marker (Willi 2018: ch.7), and by reduplication characterising perfect stems (pf. ἐκβέβηκα *ek-be-bēka*). Therefore, while escaping an interpretation as MWEs, classical Greek pre-verbed verbs represent a further stage in the evolution of VPCs from complex predicates to complex verbs (Booij and van Kemenade 2003), i.e. verbs whose meaning is still perceived as produced by two (or more) constituting elements, even though the possibility of identifying the exact contribution of each of them varies.

1.2 Objectives and theoretical framework

Much debate has surrounded the criteria for establishing if, in VPCs, the particle is part of the predicate or an independent head. The univerbation of classical Greek pre-verbed verbs apparently ensures the inclusion of the meaning of the particle into the predicate (Section 3.2), allowing us to address the problem from a different perspective and rather defining a range of effects triggered by such inclusion (Section 3.3) at the morphosyntax-semantics interface (van Valin 2005).

The theoretical framework adopted here is the lexical decomposition of predicates in event-structure patterns (van Valin and LaPolla 1997: §3.2; Rappaport Hovav and Levin 1998; Levin and Rappaport Hovav 2005). This representation highlights how the varying syntax of verb phrases reflects the distinguishing semantic trait of the lexical meaning of their head, the constant (Grimshaw 2005: ch. 2), and the relative *Aktionsart* (Vendler 1957; Dowty 1979). Therefore, it efficiently illustrates how the semantico-syntactic effects triggered by adding a particle to the predicate also depend on the meaning of the verbal component.

In lexical decomposition, a first distinction is drawn between verbs whose constant (in angle brackets) identifies a state (Section 1.2.1) or a specific way of acting (manner verbs (Section 1.2.2)).

1.2.1 (Change-of-)state verbs

When verbs involving a state predicate are static, this state is predicated as a condition of the subject.

I. States

[$x_{<STATE>}$]

e.g. *stink*

This apartment stinks

[$apartment_{<STINK>}$]

If they are dynamic, as represented by the BECOME operator, the relevant *change of* state is predicated, (II), of the subject or, (III), of the object, as the result of the action of the subject.

II. Intransitive achievements/accomplishments = unaccusatives/anticausatives (Perlmutter 1978; Haspelmath 1993)

[BECOME [$x_{<STATE>}$]]

e.g. *break* [intr.]; *melt* [intr.], *fall*

The ice melted

[BECOME [$ice_{<MELTED>}$]]

III. Causative achievements/accomplishments

[[x ACT] CAUSE [BECOME [$y_{<STATE>}$]]]

e.g. *break* [tr.]; *melt* [tr.]; *expel* [tr.]

The polar bear broke the ice

[[polar bear ACT] CAUSE [BECOME [$ice_{<BROKEN>}$]]]

The change of state may also consist in a change of position (*fall*, *expel*), involving a Theme rather than a Patient.[2]

1.2.2 Manner verbs

The constant of manner verbs predicates how the subject acts, also as a manner of movement (*run*). It is thus coded in the subevent representing the agentive role of the subject.

IV. Intransitive activities = unergatives

[x $ACT_{<MANNER>}$]

e.g. *run, eat* [intr.], *laugh, wipe* [intr.]

Tom ate (during the entire lesson)

[Tom $ACT_{<EAT>}$]

Tom laughed (during the entire dinner)

[Tom $ACT_{<LAUGH>}$]

This way of acting may secondarily entail a change of state and/or position (*wipe* [tr.]) on another entity. Therefore, these verbs enter the so-called activity/accomplishment alternation (van Valin 2013: §4.2.1), between transitive and intransitive uses. Differently from causative alterations, (II) and (III), however, the subject remains the external causer in both variants. These verbs can also be distinguished by the object they may take. Some involve Patient objects (*eat*), while others involve cognate objects of eventive value, doubling the constant of the verb in nominal form (*laugh*) and forcing a reinterpretation of the subject as producing a specific event rather than simply performing an action (see further Squeri 2026).

V. Transitive active accomplishments

[[x $ACT_{<MANNER>}$] & CAUSE [BECOME [$y_{<STATE>}$]]]

e.g. *eat* [tr.]; *laugh* [tr.]

Tom ate a bag of chips

2 The classification as Theme or Patient has the purpose of distinguishing between undergoers involved in either a change of position or a change of state (Levin–Rappaport Hovav 2005: §2.4).

[[Tom $ACT_{<EAT>}$] & CAUSE [BECOME [bag of $chips_{<EATEN>}$]]]

Tom laughed a scornful laugh

[[Tom $ACT_{<LAUGH>}$] & CAUSE [BECOME [$laugh_{<EXISTENT>}$]]]

These verbs lexically determine the way in which the subject acts to entail a change of state or produce a specific product. Therefore, the type of change of state the object undergoes or the type of produced event is bound by the verb meaning. This guarantees the consequentiality of the subevents, represented by "&" (van Valin and LaPolla 1997: 109). Weren't these blocks logically implicated, we would be facing two separate actions, represented by two verb phrases.

2. Defining the syntax and semantics of verb-particle constructions

The criteria for labelling VPCs vary especially in terms of inclusiveness (Dehé 2002: §1.2; Thim 2012: ch. 4; Luo 2019: §1.2), as results from the large array of theoretical frameworks applied to their study (Luo 2019: §2.1–2) and from the variety of languages analysed. Differently from phrasal verbs, for instance, German separable verbs allow for the particle to be both univerbated and separated from the verb. To start from a language-unspecific definition, this chapter adopts the one elaborated by the PARSEME initiative (Savary et al. 2018), which studies VPCs as a category of verbal multi-word expressions (VMWEs henceforth), annotated in a corpus of more than thirty varieties. VMWEs are predicates surfacing as (more or less regular) verb phrases rather than as single verbal lexemes, for example, *to take a nap*, *to take the bullet*, and *to take off*. They can thus be considered the lexical representation of complex predicates, "predicates which are multi-headed; they are composed of more than one grammatical element (either morphemes or words), each of which contributes part of the information ordinarily associated with a head" (Alsina, Bresnan and Sells 1997: 1; see also Amberber, Baker and Harvey 2010; Nash and Samveilan 2016).

2.1 Criteria for identifying verb-particle constructions

PARSEME's definition of VPCs combines two key characteristics: (a) the combination of verb and particle in forming a single predicate, and, (b) the idiomaticity of the resulting meaning (Dehé 2002; Thim 2012; Luo 2019).

As for criterion (a), the guidelines insist on the need for the particle to be part of the VMWE rather than the head of an autonomous prepositional phrase: verbs and particles would not be annotated in (3), (4) and (5).

(3) Tom sang (loudly) in his room (for the entire afternoon).

(4) Tom composed a song in his room.

(5) Tom took his hat off his head.

In (3) and (4) the preposition introduces an adjunct, and, thus, it is not part of the internal representation of the predicate (van Valin and LaPolla 1997: §4.4.1.1).

(3a) **In his room** [Tom ACT$_{\text{<SING>}}$]

(4a) **In his room** [[Tom ACT] CAUSE [BECOME [song$_{\text{<EXISTENT>}}$]]]

In (5), however, the preposition introduces an argument, and, therefore, is a constituting element of the predicated event (Jackendoff 2002: §4).

(5a) [[Tom ACT] CAUSE [BECOME [hat$_{\text{<OFF-HIS-HEAD>}}$]]]

In (3) and (4) the full prepositional phrase can be omitted, but the particle cannot be left in the sentence. In (5), on the other hand, one can omit the phrase, provided that the particle is left, showing that it is part of the predicate, (6).

(6) Tom took his hat off.

The particle (not) taking up a role in the predicate does not depend on the (in)transitivity of the verb: both *compose* and *put* are transitive, with a shared event template. What is relevant is the lexical meaning of the verb: *compose* predicates a change of state from non-existence to existence, while *put* predicates a change of position, whose characteristics need to be specified. The content introduced by the particle does not represent the element undergoing the predicated change of position but rather qualifies it, either alone or by heading an argument.

The fact that, once its argument is omitted, the particle can raise to direct adjacency to the verb, (7), demonstrates its inclusion in the predicate (see Dehé 2002: §1.1 and ch. 4–5; Luo 2019: ch. 4 on the unstable word order of English VPCs and the pragmatic reasons triggering the different configurations).

(7) Tom took off his hat.

When separated from the verb, the particle maintains its predicative position, following the element of which the state it represents is predicated (Los et al. 2012: §2.4). In (6), *off* occupies the same position as *red* in *we painted the room red*.

It remains to be established if verb and particle combine in a compound or a phrase (Goldberg 2016). English is a (partially) satellite-framed language, i.e. a language in which the direction of movements predicated by verbs tends to be expressed externally from the verbal lexeme (Talmy 2007: §2.6), contrarily to what happens in verb-framed languages, like Romance ones. Italian would render (5) as (8) and (6) and (7) as (9).

(8) Tommaso si è tolto il cappello dalla testa.

(9) Tommaso si è tolto il cappello.

Take off surfaces as a phrase, but, semantically, is equivalent to the Italian lexeme *togliere*. *Off* having an autonomous lexical representation, though, also allows for its use as the head of the argument it governs, whereas Italian needs to introduce it with a new preposition (*dalla*).

Structures like *take off*, therefore, cannot be excluded from VPCs because they do not meet the conditions for being classified as VMWEs. Their exclusion follows from the compositionality of their meaning: they do not meet criterion (b), establishing for VPCs a (partial) idiosyncrasy in meaning (Luo 2019, 7). This is why PARSEME relabelled VPCs as I(diomatic)VPCs. (10) and (11) would qualify as such, idiomaticity being guaranteed by the impossibility of the particle to head a prepositional phrase.

(10) Tom turned down the offer/turned the offer down.

(*Tom turned the offer down the drain.)

(11) Tom ate up the cake/ate the cake up.

(*Tom ate the cake up his nose.)

PARSEME further distinguishes between cases, like (10), in which idiomaticity is so strong that one cannot compute what element of the VMWE contributes what part of the overall meaning (IVPC.full), and cases, like (11), in which one can still identify what part of the overall meaning is contributed by the particle (IVPC.semi), even though different from the meaning the particle would have outside the IVPC.

In IVPCs.semi, the verbal head does not shift from its original meaning, and the particle alone loses its locative value. Therefore, one can still distribute the meaning represented by the lexical elements of IVPCs.semi between the blocks of the event. It suffices to allow a different meaning to the particle, specifically a telicising one (*up* = completely). This is why IVPCs.semi are also labelled as aspectual VPCs (Dehé 2002: §1.2; Luo 2019: §1.3): the type of change of state is still identified by the lexical meaning of the verb, to maintain consequentiality, but the particle insists on its completeness.

(11a) [[Tom $ACT_{<EAT>}$] & CAUSE [BECOME **up** [the $cake_{<EATEN>}$]]]

The increase in telicity triggered by the particle also affects the transitivity of the verb: *eat* can also occur intransitively, with a null object (*Tom already ate*; Section 1.2.2). This is impossible for *eat up*, since the particle needs to predicate over something.

On the other hand, IVPCs.full could only be represented by encoding the meaning represented by the full VMWE in a single portion of the event, (10a):

(10a) [[Tom ACT] CAUSE [BECOME [the offer$_{<\text{TURNED-DOWN}>}$]

2.2 Issues in identifying verb-particle constructions: A grayscale

Between the two poles represented by (3) and (4), in which the particle is not part of the predicate and only acts as the head of a prepositional phrase, and (10) and (11), in which the particle is part of the predicate and cannot head a prepositional phrase, one finds many other options. Some involve cases, like (5), whose exclusion from VPCs can at least be justified by the fact that the predicated change of position is fully identified by the particle, given the underspecification of the verbal head. Semantically, predication is thus not shared by the verb and particle but headed by the particle alone. One cannot measure the difference in meaning or telicity between the structure with and without the particle, since the verb *needs* the particle to project correctly. This section discusses other options under the provisional label of verb-particle combination. Equivalent examples in classical Greek are provided in Section 3.3.

2.2.1 Locative telicising particles with manner-of-motion verbs

The combination with a particle can turn unergative manner verbs expressing motion into unaccusative verbs of change of position. It does so by profiling an endpoint of the movement. Endpoint refers here to the point in space which measures the end of the predicated action, it might be the Target of an allative movement, (13), or the starting point of an ablative movement, (14).

(12) Tom ran for three hours.

(12a) For three hours [Tom ACT$_{<\text{RUN}>}$].

(13) Tom ran to the store in three hours.

(13a) [[Tom ACT$_{<\text{RUN}>}$] & in three hours [BECOME [Tom$_{<\text{AT-THE-STORE}>}$]]].

If the particle can predicate alone an identifiable position (see Section 3.3.5 for ablative particles being more effective in this respect), it need not to be combined with lexical content and builds a more compact unit with the verb, (14).

(14) The burglar ran off in a few seconds.

(14a) [[The burglar ACT$_{<\text{RUN}>}$] & in a few seconds [BECOME [burglar$_{<\text{OFF}>}$]]].

Differently from *eat up* (11), the particle telicises the predicate by keeping its locative meaning, either by heading a prepositional phrase or autonomously. Differently from *take off* (5) the verb could project without the particle. Nevertheless, the telicising effect of the particle can only be justified by it entering the predicate (Borer 1994), as in (11) and (5).

2.2.2 Locative telicising particles with (non-motional) manner verbs

Particles telicise manner verbs which do not only express motion. By combining with the particle, however, the manner of acting of the subject is forced into causing a change of position, (15).

(15) Tom wiped the crumbs off the table.

(15a) [[Tom $ACT_{<WIPE>}$] & CAUSE [BECOME [$crumbs_{<OFF\text{-}THE\text{-}TABLE>}$]]].

The particle can also be deprived of its argument and move to verb adjacency, (16).

(16) Tom wiped the crumbs off/Tom wiped off the crumbs.

(16a) [[Tom $ACT_{<WIPE>}$] & CAUSE [BECOME [$crumbs_{<OFF>}$]]].

The particle, thus, telicises the action, as in *eat up*, but keeps its locative meaning and can (but does not have to) head a prepositional phrase (see also Luo 2019: 81).

2.2.3 Locative telicising particles with manner verbs in locative alternations

Particles may shift from locative to telicising meaning if combined with verbs involved in locative alternations. In (15), the particle heads an argument position. This argument becomes a constituting member of the event, affected by the predicated action. (15a) should thus be more thoroughly presented as (15b).

(15b) [[Tom $ACT_{<WIPE>}$] & CAUSE [BECOME [$crumbs_{<OFF\text{-}THE\text{-}TABLE>}$]] & CAUSE [BECOME [$table_{<WITHOUT\text{-}CRUMBS>}$]]].

Being involved in the outcome of the predicated action, the endpoint can raise to the object position, in a locative alternation (Levin 1993: §2.3; van Valin and LaPolla 1997: §7.3.2.1; McIntyre 2015), (17).

(17) Tom wiped off the table.

In this restructuring, however, the particle *off* loses its locative value in favour of a telicising one, as *up* does in *eat up*. If we stick to the criterion of idiomaticity, paradoxically, the same verb-particle combination should fall within or without the category of VPCs based on the section of the same event brought in focus by the alternation. The shift in meaning, however, is provoked by the alternation of the verb-particle combination as a whole, as it happens with simplex verbs (cf. *load the truck with hay* in which *load* ≈ *fill* and *load hay on the truck*, in which *load* ≈ *put on*), and not by a shift in the semantic intake of the particle.

Such alternations are not possible for telic verbs as *put off* in 3 (**Tom put off his head [of the hat]*). With these verbs, the particle (and the lexical content

it may head) must be encoded within the change of state undergone by the argument and does not project an additional argument position whose content could compete for the direct-object slot. The head verb is not a manner verb: it only projects a change of state and the particle is thus included in that change of state. Particles telicising manner-of-motion verbs (Section 2.2.1), on the other hand, theoretically introduce a new argument slot. Differently from the alternating verbs here analysed, though, such verbs are intransitive unergatives turning into intransitive unaccusatives. Therefore, the argument coded in the prepositional phrase could not be raised to the object slot, but rather to the subject position, the key characteristic of unaccusatives being the encoding of the Patient as a surface subject, rather than as an object. A sentence like (18), in which the particle has a full localistic value, can be analysed as in (18a).

(18) Gas was running out of the tank.

(18a) [Gas $ACT_{<RUN>}$] & [BECOME [$gas_{<OUT\text{-}OF\text{-}THE\text{-}TANK>}$]] & [BECOME [$tank_{<WITHOUT\text{-}GAS>}$]]].

If we raise the content of the prepositional phrase to the argument position, as *table* in (17), but represent it as the subject of an unaccusative rather than as the object of a transitive verb, we get the idiomatic meaning of *run out* in (19) (McIntyre 2015).

(19) The tank was running out of gas.

Again, the (non)-idiomatic meaning of a verb-particle combination depends on the section of the event brought in focus by argument selection.

2.2.4 Prepositional verbs

There exist verbs whose only argument is introduced by a particle, as *look at* (*Tom looked at Mary*), as well as verbs which can draw a second participant into the action by combining with a particle, as *laugh at* (*Tom laughed at Mary*). The particle maintains its localistic value and cannot be separated from the verb (**Tom looked/laughed Mary at*, see further Thim 2012: 2.3.2). The argument is not understood as undergoing any change of state which can be predicated by the particle, it simply represents the Target. The lexical meaning of the verb is, thus, low in transitivity and telicity (Hopper and Thompson 1980). Nevertheless, the particle draws another participant into the action (**Tom looked/laughed Mary*) and does not head an adjunct with locative meaning (*Tom laughed loudly at the party*), it can thus be considered part of the predicate. PARSEME proposes a tentative category of inherently adpositional verbs (IAV henceforth) for verbs whose combination with a particle is either always mandatory or mandatory if one aims at producing a specific meaning (cf. *come* and *come across*). However, verbs fall under this definition only if they bear an idiomatic meaning, unlike *look at* and *laugh at*.

2.2.5 Secondary idiomatic meanings

Verb-particle combinations with compositional value can develop secondary idiomatic meanings. The degree of idiomaticity can vary and is triggered by different factors. The verb phrase *take out*, which is fully compositional in *Tom took out the rubbish*, has a particular nuance in *Tom would love to take Mary out*.

A specific lexical content can become so standardised for the prepositional phrase headed by the particle of a verb-particle combination that it may be omitted, in a semantic specialisation (Squeri 2023). The meaning of *take off* in (21) and (21a) could be interpreted as the anticausative version of (20) and (20a), in which the content of the prepositional phrase is omitted.

(20) The commander took the plane off land at 6 AM.

(20a) at 6 AM [[The commander ACT] CAUSE [BECOME [plane $_{\text{<OFF-LAND>}}$]]].

(21) The plane took off (land) at 6 AM.

(21a) at 6 AM [BECOME [plane $_{\text{<OFF-(LAND)>}}$]]].

Finally, the idiomatic meaning of verb-particle combinations can vary due to the varying content of their arguments (Dixon 2022: §1.3).

(22) Tom looked up the chimney.

(23) Tom looked up the word.

The meaning of *chimney* allows for *up* to combine with it in a prepositional phrase and maintain its the localistic value. The meaning of *word* disallows the combination with *up* in a prepositional phrase. *Word* is thus interpreted as the object of the idiomatic VPC *look up*.

2.2.6 Summary

Idiomaticity grants the interpretation of verb-particle combinations as VMWEs, since the unpredictable semantic intake of the particle prevents from assessing the pair as a standard combination at phrase level. The opposite, however, is not true: combinations which are not idiomatic can still be interpreted as VPCs. Idiomaticity is thus one parameter among others, and, as observed in this section, is more a graded than an exclusive feature (Luo 2019: 127). In other instances, the particle needs to be taken as part of the predicate since it represents a constituting element of its event structure, even when keeping its localistic value (Goldberg 2016). By paralleling the variables analysed in this section with classical Greek pre-verbed verbs, the next section shows how the degree of idiomaticity of VPCs correlates to the need of accommodating the particle in the event structure headed by the base verb, by keeping the necessary consequential link with its lexical meaning.

3. Particle-induced changes in the morphosyntax and semantics of classical Greek pre-verbed verbs

3.1 A working hypothesis

The idea of deriving the idiomatic meaning of VPCs from the one produced by combining the localistic meaning of the particle with the lexical meaning of the verb is not new, and has become widespread after the recognition of the human tendency to reconceptualise non-spatial aspects of reality in spatial terms (Lakoff and Johnson 1980; Luo 2019: §2.1.2 for an overview).

This chapter rather insists on how the addition of the meaning represented by an originally localistic particle to a simple predicate might shift its event template to one which is typical—but not exclusive—of predicates expressing a change of position (Luo 2019: ch. 3; Goldberg 2016). Non-locative meanings of VPCs would not be derived from locative ones via a purely metaphorical shift. The activation of either a locative or non-locative meaning would rather be linked to how the new portion of meaning represented by the particle is accommodated into the event template activated by the lexical meaning of the verb. The high degree of variability of lexical meanings of verbs will then better account for the grayscale between full compositionality and full idiomaticity observed in Section 2.2 for VPCs.

We assume that, as observed in Sections 2.1 and 2.2, when the particle enters the predicate, its meaning is accommodated into the event either within an already available change-of-state subevent, or by being assigned to a new one, whose presence becomes mandatory (hence the mandatory transitive use of *eat up* with respect to *eat*, (11) and (11a). In both cases, the type of change of state headed by the particle must adapt to the lexemic content provided by the meaning of the verb: with change-of-state verbs (Section 1.2.1), the type of change-of-state headed by the particle must combine with the type of change-of-state already identified by the verb meaning. With manner verbs, it must follow from the way of acting performed by their subject. If there were not a consequential link between the subevents headed by the verb and the particle, they could not be thought of as constituting blocks of the same event, and monoclausality, a characterising feature of complex predicates (Baker and Harvey 2010, Butt 2010), would be missing. A bi-clausal interpretation might be perceived as impossible, since the structure has a single verbal head, but some interpretations of VPCs have supposed that they actually involve a small clause (for example, *Tom saw [Mary play the piano]*) with an unrepresented verbal head (Dehé 2002: §2.2; Goldberg 2016: §3.4). The difference between small clauses and subevents is that the arguments of subevents are governed by the main verb, while the arguments of small clauses are governed by the predicative element in the small clause.

3.2 The lexical status of pre-verbed verbs

The hypothesis put forward here (Section 3.1) rests on three assumptions: (A) In pre-verbed verbs, particles enter the predicate and, thus, (B) a bi-clausal interpretation is prevented, and, (C) when particles enter the predicate, the projected event structure needs to involve a (change of) state. Assumption (C) has been proven by Tenny (1994: §1.2.7 for modern languages) and Romagno (2022 for pre-modern Greek and Latin), but, along with (B), it depends on (A) being validated.

By instantiating in a single verbal lexeme, classical Greek pre-verbed verbs strongly testify for (A), and exclude the phrase interpretation in favour of the compound one. According to Pompei (2014), the evolution from the free co-occurrence of verbs and localistic adverbs to compounds stored in the lexicon—through an intermediate stage of complex verbs—is already accomplished in Classical Greek, previous stages being testified by Homer. This is why classical Greek texts were chosen to build the dataset of the examples provided in Section 3.3, with a strong preference for prose, less prone to the tendency to imitate Homeric diction.

One piece of proof of pre-verbed verbs having raised to word status is that they serve as the base for further preverbation, for example, προκαταλέγω *prokatalegō* ← [πρό + καταλέγω *pro* + *katalegō*] (Zanchi 2019), and derivation, for example, κατάλογος *katalogos* ← καταλέγω *katalegō*, and they bear a single stress (Giannakis 2023). Horrocks and Stavrou (2007) have argued that, having a grammaticalised way of predefining aspect via stem selection, "ancient" Greek prevents telicity or perfectivity (see Bertinetto and Delfitto 2000 on the differences between the two; Rodríguez Adrados 1992: 442–449 on the relationship between stem selection and preverbation in "ancient" Greek) to be further triggered at sentence-level. The telicisation provoked on *run* by its combination with *to the store* (Section 2.2.1) cannot be triggered on Greek verbs by the same means. It can, however, be triggered by preverbation (Section 3.3.2), a fact which confirms that the preverb is a constituting element of the verb, which is only afterwards marked for aspect.[3]

Other data, however, support the interpretation of classical Greek pre-verbed verbs as early forms of univerbation, occupying an intermediate stage of phrasal compounds before reaching the stage of fully-fledged compounds. Phrasal compounds still display, in synchrony, recognisable internal structuring respecting the rules of the productive syntax, but can also be treated as standard lexical entries (Lehmann 2020), and act as productive bases for further preverbation and derivation. The perception of the constituting elements of pre-verbed verbs as partially autonomous is shown by (A) the fact that morphological material marking tense or aspect, i.e. the augment and reduplication, can still

3 Preverbation impacting on *Aktionsart*, a lexical feature (van Valin 2005: 32), and not on aspect, a grammatical feature, also secures the Greek–English parallel here proposed from any bias stemming from the different marking of aspect displayed by these two languages.

intervene between the two, and by (B) the fact that the preverb can still govern case (but see Section 3.3.1). In (24), the use of the genitive cannot be explained by its inherent ablative meaning (Luraghi 2003: §2.2.2.1), since the Samians are the Target rather than the starting point of the action.

(24) Thucydides, *Histories*, 1.115.2

κατεβόων	τῶν	Σαμίων
kateboōn	*tōn*	*Samiōn*
yell-at.impf.ind.act.3pl	the.gen.pl.m	Samian.gen.pl.m

"They yelled at the Samians."

Classical Greek pre-verbed verbs thus represent a special case in which the constituting elements of what is semantically perceived as a lexeme are partially autonomous and functional morphemes at syntax level. This allows us to interpret pre-verbed verbs as complex predicates and, only functionally, as VMWEs.

This assumption has some noteworthy entailments:

1. Preverbs are not simple prefixes. This aligns with Indo-European words being resistant to prefixation (Pinault 1995: §2) and classical Greek verbs never being otherwise prefixed by elements which can autonomously function as adverbs (forms like εὐλογέω *eulogeō* "to praise" are denominals/deadjectivals drawn from compounds built with adverbs [εὔλογος *eulogos* ← εὖ *eu* "well" + λόγος *logos* "speech"], see Duhoux 2000: §24). Preverbs are words with two alternative sentence positions: heads of prepositional phrases or verb-adjacent lexemes. When verb-adjacent, they do not simply bring a semantic addition to the verb, but head their own argument position and, thus, represent an independent subevent (Rappaport Hovav and Levin 1998), with the relevant "selectional" restrictions (Resnik 1996), as they would do when heading prepositional phrases. When that position is not saturated, (16), as in structures with null objects licensed by the restrictions provided by the meaning of the head verb (Cennamo 2017), the restriction provided by the preverbs is taken as a sufficient semantic information, and the valency of the predicate is reduced.
2. The fact that particles represent a subevent explains why they can telicise the predicate in which they are included even though they do not display any telicising meaning when used as adverbs or as heads of prepositional phrases, for example, there is no autonomous use of ἀπό *apo* which has the meaning of "completely" it takes up in ἀποθνῄσκω *apo-thnēskō* "to die" (see Luraghi 2003: §3.1; Bortone 2010: ch. 2; Beekes 2011: §17.5.1) This would be paradoxical if we understood preverbation as a simple semantic addition to the verb meaning.

3. The maintenance of the perception of the word boundary between preverb and verb might have helped the reanalysis of preverbs as pure heads of change-of-state subevents (Section 3.3.5). This function might have been attributed to them in analogy with the one they have when they truly represent an endpoint in space, if combined with verbs expressing a change of position or a manner of motion.

3.3 From endpoints to telicity-markers: How preverbs interact with verb meanings

The present section analyses how the varying impact of preverbation can be connected with two interrelated aspects: the internal structuring of the event projected by the verb and the distinguishing semantic trait of its lexical meaning, with special reference to its capacity to express—or be forced to express—motion. The relevance of the lexical meaning of verbs on their belonging to the core or to the periphery of the classes in which they are organised with respect to *Aktionsart* and transitivity (Section 1.2) has indeed been shown for modern languages (Sorace 2000).

The different effects of preverbation here analysed for classical Greek align with the ones observed for verb-particle combinations in English candidates for a VPC interpretation (Section 2.2).

3.3.1 Change-of-position verbs: Preverbs as inherent endpoints in space

Dynamic change-of-position verbs involve unaccusatives, predicating a change of position of the subject, and causatives, predicating a change of position provoked by the subject on the object (see Section 2.1). The reaching of this new position marks the end of the action.

In both groups one finds satellite-framed and verb-framed predicates (Pinault 1995: §7; Pompei 2014: §4.2), see Table 5.1.

Verbs of change of position	Satellite-framed	Verb-framed
Intransitive (unaccusative)	ἔρχομαι *erkhomai* “come”	πίπτω *piptō* “fall”
Transitive (causative)	τίθημι *tithēmi* “put”	ὠθέω *ōtheō* “push”

Table 5.1 Types of change-of-position verbs attested in classical Greek.

Satellite-framed verbs do not autonomously profile the position to be reached, so that this information *must* be provided by a prepositional phrase, if it is not pragmatically inferable (e.g. *They came [scil. here] at night*).

(25) Thucydides, *Histories*, 1.26.3

ἦλθον	γὰρ	ἐς	τὴν	Κέρκυραν
ēlthon	*gar*	*es*	*tēn*	*Kerkyran*
come.aor.ind.act.3pl	prt	to	the.acc.sg.f	Corcyra.acc.sg.f

οἱ	[...]	φυγάδες
hoi	[...]	*phugades*
the.nom.pl.m	[...]	exiled.nom.pl.m

"The exiled [...] had come to Corcyra." (on this meaning of ἐς *es*, see Luraghi 2003: §3.3)

(25a) [BECOME [exiled$_{\text{<IN-CORCYRA>}}$]]

(26) Xenophon, *Anabasis*, 7.1.22

θέσθε	τὰ	ὅπλα	ἐν
thesthe	*ta*	*hopla*	*en*
put.aor.imp.act.2pl	the.acc.pl.n	weapons.acc.pl.n	in

τάξει
taxei
line.dat.sing.f

"Put your weapons in line."

(26a) [[You ACT] CAUSE [BECOME [weapons$_{\text{<IN-LINE>}}$]]]

When these verbs are pre-verbed, the preverb becomes the fixed representation of the portion of meaning characterising the change-of-state subevent. It may alone identify the endpoint of motion or recall anaphorically backgrounded information (Horrocks 1981: §I.B.3).

(27) Demosthenes, *Speech*, 34.16

οὐκ	ἔφασκεν	οὔτε	τὰ	χρήματα
ouk	*ephasken*	*oute*	*ta*	*khrēmata*
not	say.impf.ind.act.3sg	not	the.acc.pl.n	goods.acc.pl.n

ἐντεθεῖσθαι	τοῦτον
entetheisthai	*touton*
put-in.prf.inf.mid	dem.acc.m.sg

"He denied that he had put the goods in (*sc.* in the ship previously mentioned)."

(27a) [[He ACT] CAUSE [BECOME [goods$_{<IN>}$]]]

The preverb might also project an argument position with the relevant "selectional" constraints, e.g. ἐν *en* must combine with a container (Luraghi 2003). In Homer and in classical Greek poetry, (28), the argument can be directly linked to the pre-verbed verb, whereas in classical Greek prose the preverb is rather repeated as the head of a prepositional phrase, (29).

(28) Euripides, *Orestes*, l. 1529

ἐξῆλθον	δόμων
exēlthon	*domōn*
come-out.aor.ind.act.1sg	house.gen.pl.m

"I came out of the house."

(29) Demosthenes, *Speech*, 27.25

ἐκ	τῆς	οἰκίας	ἐξῆλθεν
ek	*tēs*	*oikias*	*exēlthen*
out	the.gen.sg.f	house.gen.sg.f	come-out.aor.ind.act.3pl

"They came out of the house."

As highlighted by the representation in angle brackets, in satellite-framed verbs the preverb is perceived as the key constituent of the verb meaning. It defines the nature of the change of position undergone by the Theme, rather than heading a second argument. These pre-verbed verbs are thus more likely to be perceived as unitary lexical items (*to go out → to exit*), depriving their particle of its syntactic function. According to Pompei (2014), the particles act as incorporated elements. This combines with the fact that, in Greek, cases also display directional value (Luraghi 2003). In (28), δόμων *domōn* "of the house" (collective noun) can also be interpreted as an ablative genitive triggered by the addition of ἐκ *ek* "out of" forcing the verb into expressing an ablative movement. As opposed to common definitions of VPCs, however, this has nothing to do with the particle being deprived of its locative value, quite the contrary.

Passage (28) should therefore be represented as (28a), while (29) is rather analysed as (29a), in which the content of the prepositional phrase is coded as an adjunct, but takes scope within the change-of-state subevent (van Valin and LaPolla 1997: §4.4.1.1–2).

(28a) [BECOME [I$_{<OUT\text{-}OF\text{-}THE\text{-}HOUSE>}$]].

(29a) [BECOME **out of the house** [they$_{<OUT>}$]].

"Verb-framed" predicates already profile the endpoint in their lexical meaning.

The verb πίπτω *piptō* "to fall" inevitably profiles a downward motion, and ὠθέω *ōtheō* "to push" a motion away from the subject, (30) and (31).

(30) Xenophon, *Hellenica*, 4.4.12

ἐν	ὀλίγῳ	πολλοὶ	ἔπεσον
en	*oligō*	*polloi*	*epeson*
in	short.dat.sg.m	many.nom.pl.m	fall.aor.ind.act.1pl.

"Many fell within a short time."

(30a) within a short time [BECOME [many$_{\text{<DOWN>}}$]].

(31) Lysias, *Speech*, 1.24

ὤσαντες	δὲ	τὴν	θύραν
ōsantes	*de*	*tēn*	*thuran*
push.aor.ptcp.act.nom.pl.m	ptc	the.acc.sg.f	door.acc.sg.f

"Having pushed the door."

(31a) [[They ACT] CAUSE [BECOME [door$_{\text{<AWAY>}}$]]].

The presence of a prepositional phrase is thus optional, and its content is accommodated as an adjunct modifying the change of state, (32) and (33).

(32) Xenophon, *Hellenica*, 3.4.24

οἱ	μὲν	[…]	ἐν	τῷ	ποταμῷ
hoi	*men*	*[…]*	*en*	*tō*	*potamō*
the.nom.pl.m	ptc	[…]	in	the.dat.sg.m	river.dat.sg.m

ἔπεσον
epeson
fall.aor.ind.act.3pl

"Some […] fell into the river."

(32a) [BECOME **into the river** [some$_{\text{<DOWN>}}$]].

(33) Plato, *Timaeus*, 79b

(*τὸ πνεῦμα)*	τὸ	πλησίον	ἐκ	τῆς
(*to pneuma*)	*to*	*plēsion*	*ek*	*tēs*
breath.nom.sg.n	the.acc.sg.n	adjacent.acc.sg.n	from	the.gen.sg.f

ἕδρας	ὠθεῖ
hedras	*ōthei*
seat.gen.sg.f	push.prs.ind.act.3sg

"Breath pushes the adjacent body from its seat."

(33a) [[Breath ACT] CAUSE [BECOME **from its seat** [adjacent-body$_{<AWAY>}$]]].

The meaning represented by the preverb in verb-framed predicates of change of position can virtually be coded both inside the verb constant (34a) and (35a), as part of the lexical meaning of the pre-verbed verb, or as an adjunct modifying the change of state already defined by the verb (34b) and (35b). What is relevant for the present argument is that the final interpretation of both representations is basically equivalent.

(34) Thucydides, *Histories*, 1.144.4

τόν	τε	βάρβαρον	ἀπεώσαντο
ton	*te*	*barbaron*	*apeōsanto*
the.acc.sg.m	and	barbarian.acc.sg.m	push-away.aor.mid.3pl

"And they pushed the barbarian away."

(34a) [[They ACT] CAUSE [BECOME [barbarian$_{<AWAY\text{-}AWAY>}$]]].

(34b) [[They ACT] CAUSE [BECOME **away** [barbarian$_{<AWAY>}$]]].

(35) Thucydides, *Histories*, 1.70.6

νικώμενοι	ἐπ'	ἐλάχιστον
nikōmenoi	*ep'*	*elakhiston*
win.prs.ptcp.pass.nom.pl.m	as	little.adj.superl.acc.sg.n

ἀναπίπτουσιν
anapiptousin
fall-back.prs.ind.act.3pl

"If beaten, they fall back as little as possible."

(35a) as little as possible [BECOME [they$_{<BACK\text{-}DOWN>}$]]].

(35b) as little as possible [BECOME **back** [they$_{<DOWN>}$]]].

Semantically the preverb might either double the endpoint already profiled by the verb meaning, (34), or add a further endpoint reached by moving in the direction profiled by the verb, (35). In the first case, the endpoint marked by the preverb is simply overabundant.

As observed for satellite-framed predicates, the pre-verbed verb can directly head an argument or the preverb can be doubled as the head of a prepositional phrase.

With change-of-position verbs, thus, the preverb gets easily accommodated within the verb meaning, both in terms of event structure, since these verbs already project a change of state, and in terms of semantics, since that change

of state is locative. The preverb is either the only element profiling the endpoint of that change, with satellite-framed predicates, or combines with the endpoint already profiled by the lexical meaning, with verb-framed predicates. It thus moves as close as possible to the predicate, even by keeping its localistic value.

Idiomatic meanings can further be derived from the meaning of the whole pre-verbed verb, (36).

(36) Plato, *Theaetetus*, 145c

μὴ	ἀναδύου	τὰ	ὡμολογημένα
mē	*anaduou*	*ta*	*hōmologēmena*
not	stand-up-from.pres.imp.mid.2sg	the.acc.pl.n	agreement.acc.pl.n

"Do not draw back from your agreement."

The partially idiomatic meaning of ἀναδύομαι *anaduomai* in (36) is not provoked by preverbation, but by a secondary meaning shift of the whole pre-verbed verb, triggered by the content of the argument (Section 2.2.5).

3.3.2 Manner-of-motion verbs: Preverbs as additional endpoints in space

The lexical meaning of unergative manner-of-motion verbs does not qualify a subevent of change of position but rather qualifies the performance of the subject. The action does not profile an endpoint and is atelic (see Section 2.2.1).

(37) Xenophon, *Anabasis*, 6.1.14

ἔπλεον	ἡμέραν	καὶ	νύκτα
epleon	*hēmeran*	*kai*	*nykta*
sail.impf.ind.act.3pl.	day.acc.sg.f	and	night.acc.sg.f

"They sailed for a day and a night."

(37a) for a day and a night [They $ACT_{<SAIL>}$].

In English, these verbs can be telicised by the addition of a prepositional phrase profiling an endpoint (Section 2.2.1). As explained in Section 3.2, in classical Greek this is not possible: the prepositional phrase would simply provide a direction for the atelic movement. However, this effect can be triggered by preverbation, which, entering the predicate, directly acts on the meaning of the verb, and, thus, on the type of event it projects.

The preverb might be the only lexical element representing the subevent, (38), but can also govern an argument, (39), or be doubled as the head of a prepositional phrase, (40).

(38) Thucydides, *Histories*, 1.65.1

ξυνεβούλευε	[…]	ἐκπλεῦσαι
synebouleue	[…]	*ekpleusai*
order.impf.ind.act.3pl	[…]	sail-out.aor.inf.act

"He ordered […] to sail out."

(38a) [[They ACT$_{<SAIL>}$] & [BECOME [they$_{<OUT>}$]]].

(39) Thucydides, *Histories*, 2.24.1

ἢν	[…]	οἱ	πολέμιοι	[…]	ἐπιπλέωσι
ēn	*[…]*	*hoi*	*polemioi*	*[…]*	*epipleōsi*
if	[…]	the.nom.sg.m	enemies.nom.sg.m	[…]	sail-at.prs.subj.act.3pl

τῇ	πόλει
tē	*polei*
the.dat.sg.f	city.dat.sg.f

"If […] the enemies […] sailed against the city."

(39a) [[enemies ACT$_{<SAIL>}$] & [BECOME [enemies$_{<AT THE CITY>}$]]].

(40) Thucydides, *Histories*, 1.52.1

ἐπέπλευσαν	ἐπὶ	τὸν	ἐν	τοῖς
epepleusan	*epi*	*ton*	*en*	*tois*
sail-to.aor.ind.act.3sg	to	the.acc.sg.m	in	the.dat.pl.n

Συβότοις	λιμένα
Sybotois	*limena*
Sybota.dat.pl.n	harbour.acc.sg.m

"They sailed against the harbour at Sybota."

(40a) [[They ACT$_{<SAIL>}$] & [BECOME [they$_{<AT THE HARBOUR>}$]]].

The change-of-state subevent cannot be headed by the simplex verb and is thus headed by the preverb alone, autonomously projecting an argument position. If that position is left unsaturated, the semantic information projected by the preverb-head in the form of "selectional" restrictions is the only element characterising that subevent, or recalling anaphorical information. Moreover, specific content can be assumed to be included in that subevent, even though left unexpressed, triggering a (partially) idiosyncratic meaning. The pre-verbed verb καταπλέω *katapleō* for instance, develops the meaning of "to sail back *home*" from the compositional one of "to sail back".

3.3.3 Manner verbs with Patient-Theme objects: Preverbs heading a second argument

At the border between change-of-state verbs and manner verbs, there are verbs predicating a specific way of acting of the subject, (41), which may also entail a change of state over another element, (42) (see Sections 2.2.2 and 2.2.3).

(41) Aristotle, *Problems*, 864a

τὰ	φάρμακα	καθαίρει
ta	*pharmaka*	*kathairei*
the.nom.pl.n	drugs.nom.pl.n	clean.prs.ind.act.3sg

"Drugs clean."

(41a) [Drugs $ACT_{<CLEAN>}$].

(42) Plato, *Timaeus*, 72c

πάντα	ἡ	σπληνὸς	καθαίρουσα
panta	*hē*	*splēnos*	*kathairousa*
all.acc.pl.n	the.nom.sg.f	spleen.gen.sg.f	clean.prs.prtc.act.nom.sg.f

αὐτὰ	δέχεται	μανότης
auta	*dekhetai*	*manotēs.*
them.acc.pl.n	absorb.prs.ind.mid.3sg	loose-texture.nom.sg.f

"The loose texture of the spleen cleanses and absorbs them all."

(42a) [[Spleen $ACT_{<CLEAN>}$] & CAUSE [BECOME [$impurities_{<CLEANED>}$]]].

The characteristics of this change of state must follow from the way of acting encoded in the verb meaning and may also involve a change of position. If they do, the object can be understood as both a Patient and a Theme, altered both in condition and position. There is no example of simple καθαίρω *kathairō* "to clean" with a prepositional phrase defining the spatial endpoint of the action in Attic prose, but this structure is frequent in the contemporary Ionic prose of the *Hippocratic Corpus*, (43).

(43) Hippocrates, *On the Nature of Women*, 23.1, p. 25, 1–2 Bourbon (7.342.2–3 L.)

ὑφ'	οὗ	αἷμα	καθαίρεται	ἐκ
uph'	*hou*	*aima*	*kathairetai*	*ek*
by	rel.gen.sg.n	blood.nom.sg.n	clean.prs.ind.pass.3sg	from

τῶν	ὑστερέων
tōn	*hystereōn*
the.gen.pl.f	uterus.gen.pl.f

"(A suppository) by which blood is cleaned from her uterus."

(43a) [from her uterus [Suppository ACT$_{<\text{CLEAN}>}$] & CAUSE [BECOME [blood$_{<\text{CLEANED}>}$]]].

The content of the prepositional phrase can take scope over the whole event, (43a): it qualifies the way in which the subject acts, and, thus, the consequential change of state undergone by the Patient.

A different interpretation is however possible, (43b), in which the endpoint is drawn into the event, as a secondary affected element involved in the chain of subevents. The Patient is reinterpreted as a Theme, whose movement allows the interaction with the element representing the endpoint. All the subevents, however, must still logically follow from the manner of acting predicated of the subject.

(43b) [[Suppository ACT$_{<\text{CLEAN}>}$] & CAUSE [BECOME [blood$_{<\text{OFF-HER-UTERUS}>}$]] & CAUSE [BECOME [uterus$_{<\text{WITHOUT-BLOOD}>}$]]]

When a preverb is added, (44), the endpoint it expresses is fixed within the event. Therefore, differently from the simplex verb, the pre-verbed verb is mandatorily transitive: a Theme is needed for the endpoint to be reached and the action to end (Tenny 1994: §1.2.7; Romagno 2022). If the argument slot headed by the preverb is saturated, an interpretation like (43b) is mandatory, (44a).

(44) Aristotle, *On Colours*, 795b

ὅταν	[…]	ἐξ	αὐτῆς	(πορφύρας)	τὴν
hotan	[…]	*ex*	*autēs*	*(porphuras)*	*tēn*
when	[…]	off	it.gen.sg.f	murex.gen.sg.f	the.acc.sg.f

ὑγρασίαν	ἐκκλύσωσι
hygrasian	*ekklusōsi*
moisture.acc.sg.f	clean-off.prs.subj.act.3pl

"When […] they have cleaned the moisture off the murex."

(44a) [[They ACT$_{<\text{CLEAN}>}$] & CAUSE [BECOME [moisture$_{<\text{OFF-THE-MUREX}>}$]] & CAUSE [BECOME [murex$_{<\text{WITHOUT-MOISTURE}>}$]]].

If the position is left unsaturated the preverb can be understood as simply adding a change of position to the change of state undergone by the Patient, (45a), or to anaphorically retrieve a more specific reference point, (45b).

(45) Aristotle, *Problems*, 860a

οὐκ	ἀποκαθαρθέντος	τοῦ	φλέγματος
ouk	*apokatharthentos*	*tou*	*phlegmatos*
not	clear-away.aor.prtc.pass.gen.sg.n	the.gen.sg.n	phlegm.gen.sg.n

"The phlegm not being cleared away (from the body)."

(45a) [[BECOME [phlegm$_{<\text{CLEARED}>}$]] & [BECOME [phlegm$_{<\text{AWAY}>}$]]].

(45b) [[BECOME [phlegm$_{<\text{OFF-THE-BODY}>}$]] & [BECOME [body$_{<\text{WITHOUT-PHLEGM}>}$]]].

The semantics of manner verbs do not allow the preverb to simply qualify the change-of-state subevent involving the object. That change of state must follow from the lexical meaning of the verb which *need not* be conceived as involving motion, (41) and (42). The endpoint represented by the preverb needs to be *added* in an additional argument slot, in order to force an interpretation of the object as *also* undergoing a change of position. The action is thus understood as involving two arguments: the Patient, reinterpreted as a Theme, and the endpoint reached by/deprived of the Theme. The fact that both are coded as arguments in autonomous subevents is shown by their competing for the direct-object slot, in a locative alternation, compare (45) with (46) (de la Villa 2017).

(46) Aristotle, *Problems*, 958b

τὸ	ἀποκεκαθάρθαι	τὰ	ὄμματα
to	*apokekatharthai*	*ta*	*ommata*
the.nom.sg.n	clean-off.prf.inf.mid	the.nom.pl.n	eyes.nom.pl.n

"The fact that the eyes are cleaned up."

(46a) [BECOME [eyes$_{<\text{CLEANED-UP}>}$]].

Most of these verbs can already alternate in the simplex form, but the addition of a preverb helps in aligning syntax with semantics, providing two semantically different types of heads for two semantically different types of arguments. However, in the alternated form, (46), the value of the preverb moves from locative, and compositional, to telicising, and idiomatic.

3.3.4 Manner verbs with eventive meaning: Preverbs heading a first argument

Many manner verbs predicate a type of performance of the subject which cannot involve a Patient or Theme. The action is identified by the type of event produced while acting: one can tell *laughing* apart from *crying* by the type of produced event (see Section 2.2.4), (47).

(47) Plato, *Euthydemus*, 303d

γελῶντες	[…]	ὀλίγου	παρετάθησαν
gelōntes	*[…]*	*oligou*	*paretathēsan*
laugh.prs.prtc.act.nom.pl.m	[…]	nearly.adv	faint.aor.ind.pass.3pl

"Laughing […] until they nearly fainted."

The verb can sometimes combine with a syntactic object, involved in the action by being the Target of such an eventive product, (48) and (48a). Given the eventiveness of the Theme, the Target is less affected than the one of Section 3.3.3 and gets more likely coded as an adjunct in a prepositional phrase, (49) and (49a).

(48) Xenophon, *Symposium*, 2.19

ἦ	τόδε	γελᾶτε;
ē	*tode*	*gelate?*
maybe	dem.acc.sg.n	laugh.prs.ind.act.2pl

"Are you laughing at this?"

(48a) [[you ACT$_{<\text{LAUGH}>}$] & CAUSE [BECOME [this$_{<\text{DERIDED}>}$]].

(49) Xenophon, *Symposium*, 2.18

ἐπ'	ἐκείνῳ	γελᾶτε
ep'	*ekeinō*	*gelate*
at	dem.dat.sg.m	laugh.prs.ind.act.2pl

"You laugh at him."

(49a) at him [you ACT$_{<\text{LAUGH}>}$].

By opening a new argument slot, thus, preverbation gives the opportunity to keep the Target in the verb phrase by aligning its semantic role with coherent morphosyntactic coding, (50).

(50) Plato, *Cratylus*, 413b

καταγελᾷ	μου	οὗτος
katagela	*mou*	*houtos*
laugh-at.prs.ind.act.3sg	I.gen.sg	dem.nom.sg.m

"This one laughs at me."

(50a) [He ACT$_{<\text{LAUGH}>}$] & CAUSE [BECOME [laugh$_{<\text{AT-ME}>}$] & CAUSE [BECOME [I$_{<\text{LAUGHED AT}>}$]].

In rare cases, the pre-verbed verb can alternate by raising the Target to object position (Romagno 2008), preferably surfacing as the subject of a passive form, (51). Cases with the accusative are more likely found in poetry, for example, Euripides, *Bacchae*, l. 286.

(51) Plato, *Theaetetus*, 175b

ὁ	τοιοῦτος	ὑπὸ	τῶν	πολλῶν
ho	*toioutos*	*hypo*	*tōn*	*pollōn*
the.nom.sg.m	dem.nom.sg.m	by	the.gen.pl.m	many.gen.pl.m

καταγελᾶται
katagelatai
deride.pres.ind.pass.3sg

"He is derided by many."

(51a) [[Many ACT$_{\text{<LAUGH>}}$] & CAUSE [BECOME [he$_{\text{<LAUGHED AT>}}$]].

Again, in the alternated form, the value of the preverb turns from localistic to telicising.

In some cases, the preverb simply intensifies the action, insisting on the complete production of the relevant event. This is more likely achieved with ablative preverbs, unapt to profile a Target: ἐκ-γελάω *ek-gelaō* simply means "to laugh loudly".

3.3.5 Change-of-state verbs, and manner verbs not implying motion

In the previous cases, preverbs are coded in a subevent of change of state which, semantically, represents an endpoint in space, but, functionally, represents the end of the action.

This ambivalence is clear in alternating verbs. In two cases, however, preverbs take up an exclusively telicising function (see Section 2.1). Firstly, when combined with change-of-state verbs without a change of position implied, as κτείνω *kteinō* "to kill", they overabundantly mark the telicity of the change of state, (52) and (53).

(52) Thucydides, *Histories*, 2.102

ἔκτεινε	τὴν	μητέρα
ekteine	*tēn*	*mētera*
kill.impf.ind.act.3sg	the.acc.sg.f	mother.acc.sg.f

"He killed his mother."

(52a) [[He ACT] CAUSE [BECOME [mother$_{\text{<DEAD>}}$]]].

(53) Isocrates, *Speech*, 9.38

Κῦρος	δὲ	τὸν	πατέρα	τὸν
Kyros	*de*	*ton*	*patera*	*ton*
Cyrus.nom.sg.m.	ptc	the.acc.sg.m	father.acc.sg.m	the.acc.sg.m

τῆς	μητρὸς	ἀπέκτεινεν
tēs	*mētros*	*apekteinen*
the.gen.sg.f	mother.gen.sg.f	slay.impf.ind.act.3sg

"Cyrus slew his mother's father."

(53a) [Cyrus ACT] CAUSE [BECOME **completely** [father$_{<DEAD>}$]]]

Secondly, when combined with manner verbs which cannot imply motion and are involved in the activity-accomplishment alternation, as ἐσθίω *esthiō* "to eat", the need to accommodate the preverb in a change-of-state subevent fixates the transitive use, (54) and (55).

(54) Xenophon, *Cyropedia*, 8.9

ἐσθίοντες	[…]	διάγουσιν
esthiontes	*[…]*	*diagousin*
eat.pres.prtc.act.nom.pl.m	[…]	keep-on.prs.ind.act.3pl

"They keep on eating."

(54a) [They ACT$_{<EAT>}$].

(55) Aristotle, *History of Animals*, 537a

καὶ	τὸ	δέλεαρ	[…]	κατεσθίουσιν
kai	*to*	*delear*	*[…]*	*katesthiousin*
even	the.acc.sg.n	bait.acc.sg.n	[…]	eat-up.prs.ind.act.3pl

"They eat up even the bait."

(55a) [They ACT$_{<EAT>}$] & CAUSE [BECOME **completely** [bait$_{<EATEN>}$].

By not blending into the verb meaning and being still identifiable as autonomous heads within complex predicates, preverbs might have been functionally re-analysed as simple telicisers, abstracting their function of endpoints from their semantic value of endpoints of motion events (Horrocks 1981: §3.E). This reanalysis might have been helped by (i) their ambiguous role with alternating verbs and (ii) their combination with verb-framed predicates of change of position (πίπτω *piptō* "to fall"), when the endpoint they profile doubles the one already profiled by the verb (ἐκπίπτω *ekpiptō* "to fall off"). In this latter case, the preverb could have been interpreted exclusively as a telicity marker to avoid redundancy, for example, "away" in (34b) shares the same position as "completely" in (53a) and (55a). The redundant effect, though, is only perceived if the preverb does not head an argument providing a more specific reference point. Pragmatically, the reference points which are more likely to be omitted,

being thematised by what precedes or by the contingent situation, are the ones representing the starting points of the movement, rather than its Target. This might explain why preverbs acquiring telicising function are mostly ablatival (ἐκ- *ek-*, ἀπο- *apo-* "from", ἀνα- *ana-* "up from", κατα- *kata-* "down from"). This reinterpretation of preverbs as purely functional telicisers also explains the idiosyncrasy of their choice and the apparent synonymity that they display when used as such, which they could not display when used as prepositions (cf. ἀπο-κτείνω *apo-kteinō* "to kill" with κατ-εσθίω *kat-esthiō* "to eat up"). A reuse of particles in composition with noun-morphemes with which they display an idiosyncratic meaning due to the compound being constructed in analogy with an existing one with a fully combinatorial meaning is also attested for nouns and adjectives (Rousseau 2016: §3.b.1–2), for example, in ὑποταρτάριος *hypo-tartarios* "*in* the Tartarus", ὑπο- *hypo-* "under" occurs with the idiosyncratic meaning of "in", its choice being analogical with its use in ὑποχθόνιος *hypo-khthonios* "*under* the earth", in which it displays a fully combinatorial value.

5. Conclusion and outlook

Being univerbated, classical Greek pre-verbed verbs elude by definition the label of VPCs, if understood as VMWEs, as in the definition of PARSEME. Univerbation of pre-verbed verbs, on the other hand, guarantees the inclusion of the particle in the predicate. All English borderline cases explored in Sections 2.1 and 2.2, excluded from the category of VPCs due to the lack of idiomaticity, can be paralleled to classical Greek pre-verbed verbs, semantically and functionally. Idiomaticity turns out to be a sufficient, but not a necessary parameter to assess the inclusion of a particle into the predicate, and, thus, classify a verb-particle combination as a VPC.

Coherently with their original localistic value, preverb-particles profile the endpoint of motion events or events whose interpretation as involving motion can be forced. All motion events expressed by pre-verbed verbs are telic since the preverb gets coded in a subevent of change of position: the action ends when the Theme changes its position according to the endpoint profiled by the preverb. When combined with verbs not involving motion, preverbs are reanalysed as telicisers, generally profiling the change of state which marks the end of an action. Therefore, in both phrasal verbs and pre-verbed verbs, the idiomaticity of the verb-particle combination, and the consequent (in)capacity of the particle to head a prepositional phrase, cannot be taken as proof of the particle (not) being part of the predicate. These aspects simply depend on the type of final state which can logically follow from the meaning of the verb combined with the particle.

The claims made here should be fully substantiated by a diachronic analysis of the function of pre-modern Greek particles as well as a typological analysis of

the outcome and function of particles of the same stem in other Indo-European languages. The telicising function of particles profiling an endpoint should also be contrasted to the one of particles profiling a path, for example, περί *peri* "all around" (see Cappelle and Declerck 2005 for English), or a non-locative information, for example, σύν *syn* "with".

Acknowledgements

This work received support from the CA21167 COST action UniDive, funded by COST (European Cooperation in Science and Technology).

Bibliography

Alsina, Alex, Joan Bresnan, and Peter Sells (eds.). 1997. *Complex Predicates.* Stanford, CA: CSLI Publications.

Amberber, Mengistu, Brett Baker, and Mark Harvey (eds.) 2010. *Complex Predicates: Cross-linguistic Perspectives on Event Structure*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511712234

Baker, Brett, and Mark Harvey. 2010. Complex Predicate Formation. In Mengistu Amberber, Brett Baker, and Mark Harvey (eds.), *Complex Predicates: Cross-linguistic Perspectives on Event Structure*, 13–47. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511712234.003

Beekes, Robert S.P. 2011. *Comparative Indo-European Linguistics. An Introduction* (Second edition revised and corrected by Michiel de Vaan). Amsterdam and Philadelphia, PN: John Benjamins. https://doi.org/10.1075/z.172

Bertinetto, Pier Marcom, and Denis Delfitto. 2000. Aspect vs. Actionality. Why They Should be Kept Apart. In Östen Dahl (ed.), *Tense and Aspect in the Languages of Europe*, 189–225. Berlin and New York: Mouton De Gruyter. https://doi.org/10.1515/9783110197099.1.189

Booij, Gert, and Ans Van Kemenade. 2003. Preverbs: An Introduction. In Geert Booij and Jaap van Marle (eds), *Yearbok of Morphology* 2003, 1–11. Dodrecht: Springer. https://doi.org/10.1007/978-1-4020-1513-7_1

Borer, Hagit. 1994. The Projection of Arguments. *University of Massachusetts Occasional Papers in Linguistics* 17(20). 19–48.

Bortone, Pietro. 2010. *Greek Prepositions from Antiquity to the Present*. Oxford and New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199556854.001.0001

Butt, Miriam. 2010. The Light Verb Jungle: Still Hacking Away. In Mengistu Amberber, Brett Baker, and Mark Harvey (eds.), *Complex Predicates: Cross-linguistic Perspectives on Event Structure*, 48–78. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511712234.004

Cennamo, Michela. 2017. Object Omission and the Semantics of Predicates in Italian in a Comparative Perspective. In Lars Hellan, Andrej Malchukov, and Michela Cennamo (eds.), *Contrastive Studies in Verbal Valency*, 251–273. Amsterdam and Philadelphia, PN: John Benjamins. https://doi.org/10.1075/la.237.08cen

Cappelle, Bert, and Renaat Declerck. 2005. Spatial and Temporal Boundedness in English Motion Events. *Journal of Pragmatics* 37. 889–917. https://doi.org/10.1016/j.pragma.2004.10.012

de la Villa, Jesús. 2017. Verbal Alternations in Ancient Greek as an Interface Between Lexicon and Syntax. In Felicia Logozzo and Paolo Poccetti (eds.), *Ancient Greek Linguistics. New Approaches, Insights, Perspectives*, 535–550. Berlin and Boston, MA: De Gruyter. https://doi.org/10.1515/9783110551754-547

Dehé, Nicole. 2002. *Particle Verbs in English. Syntax, Information Structure and Intonation.* Amsterdam and Philadelphia, PN: John Benjamins. https://doi.org/10.1075/la.59

Dixon, Robert M. W. 2022. *English Prepositions. Their Meanings and Uses*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198868682.001.0001

Dowty, David R. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montagues PTQ.* Dordrecht: D. Reidel.

Duhoux, Yves. 2000. *Le verbe grec ancien. Éléments de moprhologie et de syntaxe historiques. Deuxième édition revue et augmentée.* Louvain-la-Neuve: Peeters.

Giannakis, Georgios K. 2023. At the Crossroads of Linguistics and Philology: The Tmesis-to-Univerbation Process in Ancient Greek. In Georgios K. Giannakis et al. (eds.), *Classical Philology and Linguistics Old Themes and New Perspectives*, 175–211. Berlin and Boston, MA: De Gruyter. https://doi.org/10.1515/9783111272887-008

Goldberg, Adele E. 2016. Tuning in to the Verb-Particle Construction in English. In Léa Nash and Pollet Samvelian (eds.), *Approaches to Complex Predicates*, 110–141. Leiden and Boston, MA: Brill. https://doi.org/10.1163/9789004307094_006

Grimshaw, Jane. 2005. *Words and Structure*. Stanford, CA: CLSI Publications.

Haspelmath, Martin. 1993. More on the Typology of Inchoative/Causative Verb Alternations. In Bernard Comrie and Mara Polinsky (eds.), *Causatives and Transitivity*, 87–120. Amsterdam and Philadelphia, PN: John Benjamins.

— 2010. Comparative Concepts and Descriptive Categories in Crosslinguistic Studies. *Language* 86(3). 663–687. https://doi.org/10.1353/lan.2010.0021

Hopper, Paul, and Sandra Thompson. 1980. Transitivity in Grammar and Discourse. *Language* 56(2). 251–299.

Horrocks, Geoffrey. 1981. *Space and Time in Homer: Prepositional and Adverbial Particles in the Greek Epic*. Salem: Ayer Company.

Horrocks, Geoffrey, and Melita Stavrou. 2007. Grammaticalized Aspect and Spatio-temporal Culmination. *Lingua* 117(4). 605–644. https://doi.org/10.1016/j.lingua.2006.03.007

Jackendoff, Ray. 2002. English Particle Constructions, the Lexicon, and the Autonomy of Syntax. In Nicole Dehé et al. (eds.), *Verb-Particle Explorations*, 67–94. Berlin and New York: Mouton De Gruyter. https://doi.org/10.1515/9783110902341.67

Lehmann, Christian. 2020. Univerbation. *Folia Linguistica Historica* 41(1). 205–252. https://doi.org/10.1515/flih-2020-0007

Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. Chicago, IL and London: University of Chicago Press.

Levin, Beth. 1993. *English Verb Classes and Alternations. A Preliminary Investigation*. Chicago, IL and London: The University of Chicago Press.

Levin, Beth, and Malka Rappaport Hovav. 2005. *Argument Realisation*. Cambridge and New York: Cambridge University Press.

Los, Bettelou et al. 2012. *Morphosyntactic Change: A Comparative Study of Particles and Prefixes.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511998447

Luo, Han. 2019. *Particle Verbs in English. A Cognitive Linguistic Perspective.* Singapore: Springer Nature. https://doi.org/10.1007/978-981-13-6854-7

Luraghi, Silvia. 2003. *On the Meaning of Prepositions and Cases. The Expression of Semantic Roles in Ancient Greek.* Amsterdam and Philadelphia, PN: John Benjamins. https://doi.org/10.1075/slcs.67

McConnell, Tom. 2025. *Chronology, Dialect, and Style in Early Greek Hexameter Poetry*. Oxford: Oxford University Press. https://doi.org/10.1093/9780198932345.001.0001

McIntyre, Andrew. 2015. Particle-verb Formation. In Peter O. Müller et al. (eds), *An International Handbook of the Languages of Europe. Vol. 1: Word-Formation*, 434–449. Berlin and Boston, MA: De Gruyter. https://doi.org/10.1515/9783110246254-025

Nash, Léa, and Pollet Samvelian (eds.). 2016. *Approaches to Complex Predicates*. Leiden and Boston, MA: Brill. https://doi.org/10.1163/9789004307094

Perlmutter, David M. 1978. Impersonal Passives and the Unaccusative Hypothesis. In Jeri Jaeger et al. (eds.), *Proceedings of the Berkeley Linguistics Society* 4. 157–189.

Pinault, Georges-Jean. 1995. Le problème du préverbe en indo-européen. In André Rousseau (ed.), *Les préverbes dans les langues d'europe*, 35–59. Villeneuve d'Ascq: Presses Universitaire du septentrion.

Pompei, Anna. 2014. Verb-Particle Constructions and Preverbs in Homeric Greek between Telicization, Incorporation and Valency Change. In Annamaria Bartolotta (ed.), *The Greek Verb: Morphology, Syntax, and Semantics*. *Proceedings of the 8th International Meeting of Greek Linguistics (Agrigentum, October 1-2, 2009)*, 253–276. Louvain-la-Neuve and Walpole, MA: Peeters.

Rappaport Hovav, Malka, and Beth Levin. 1998. Building Verb Meanings. In Miriam Butt and Wilhelm Geuder (eds.), *The Projection of Arguments: Lexical and Compositional Factors*, 97–134. Stanford, CA: CLSI Publications.

Resnik, Paul. 1996. Selectional Constraints: An Information-theoretic Model and Its Computational Realization. *Cognition* 61. 127–159. https://doi.org/10.1016/s0010-0277(96)00722-6

Rodríguez Adrados, Francisco. 1992. *Nueva sintaxis del Griego Antiguo*. Madrid: Gredos.

Romagno, Domenica. 2008. Applicative and Causative: Some Further Reflections upon Verbal Prefixation in Greek and Latin. *Archivio glottologico italiano* 93(1). 80–88.

— 2022. Aspects of the Verbal Domain in Greek and Latin: Changing Valency and Actionality. In Domenica Romagno et al. (eds.), *Variation, Contact, and Reconstruction in the Ancient Indo-European Languages*, 290–332. Leiden and Boston, MA: Brill. https://doi.org/10.1163/9789004508873_013

Rousseau, Nathalie. 2016. *Du syntagme au lexique. Sur la composition en grec ancien.* Paris: Les belles lettres.

Savary, Agata et al. 2018. PARSEME Multilingual Corpus of Verbal Multiword Expressions. In Stella Markantonatou, Carlos Ramisch, Agata Savary, and Veronika Vincze (eds.),

Multiword Expressions at Length and in Depth. Extended Papers from the MWE 2017 *Workshop*, 87–147. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.1471591

Sorace, Antonella. 2000. Gradients in Auxiliary Selection with Intransitive Verbs. *Language* 76(4). 859–890. https://doi.org/10.2307/417202

Squeri, Elena. 2023. La specializzazione semantica nel sistema verbale del greco antico e nella lingua della ginecologia ippocratica. PhD thesis (Università di Genova—Sorbonne Université).

— 2026. Ancient Greek Cognate Object Constructions as an Instance of Complex Predication: Issues in (In)transitivity. *Studi e saggi linguistici* 64(1). 117–157.

Talmy, Leonard 2007. Lexical Typologies. In Timothy Shopen (ed.), *Language Typology and Syntactic Description.* 3: *Grammatical Categories and the Lexicon*, 66–168. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511618437.002

Tenny, Carol L. 1994. A*spectual Roles and the Syntax-Semantics Interface.* Dodrecht: Springer.

Thim, Stefan. 2012. *Phrasal Verbs. The English Verb-Particle Construction and its History*. Berlin and Boston, MA: De Gruyter. https://doi.org/10.1515/9783110257038

van Valin, Robert D. Jr. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge and New York: Cambridge University Press.

— 2013. Lexical Representation, Co-composition, and Linking Syntax and Semantics. In James Pustejovsky et al. (eds), *Advances in Generative Lexicon Theory*, 67–107. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-5189-7_4

van Valin, Robert D. Jr., and Randy LaPolla. 1997. *Syntax. Structure, Meaning and Function*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139166799

Vendler, Zeno. 1957. Verbs and Times. *The Philosophical Review* 66. 143–160. https://doi.org/10.2307/2182371

Willi, Andreas. 2018. *Origins of the Greek Verb*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108164207

Zanchi, Chiara. 2019. *Multiple Preverbs in Ancient Indo-European Languages. A Comparative Study on Vedic, Homeric Greek, Old Church Slavic, and Old Irish*. Tübingen: Narr Francke Attempto Verlag. https://doi.org/10.24053/9783823392743

6. The pragmatic development of ποίησον ἀγάπην *poíēson agápēn* as an intersubjective marker in post-classical Greek (4th to 6th c. CE)

Lucía Madrigal Acero

1. Introduction

Languages employ a range of mitigation strategies (Brown and Levinson 1987; Caffi 2005). From politeness markers to indirect speech, to forms of address, each speaker, language, and culture has its own way of maintaining social harmony, addressing uncertainty, or managing communicative contexts. In this chapter, I address one of the linguistic formulae that pre-modern Greek developed for attenuating interrogative and directive speech acts: ποίησον ἀγάπην *poíēson agápēn* "(*literally*) do [me] a favour > please."[1]

Cross-linguistically, there are several different sources for phrases meaning *please*, that is, phrases that are typically used to downplay commands or direct questions. Where some languages choose to use verbs of begging or asking in the first person (Latin *precor* > Italian *prego*, Standard Modern Greek (SMG henceforth) παρακαλώ *parakaló,* German *bitte*), others prefer to address the interlocutor more directly with the use of an imperative form (Italian *scusi*, French *excusez-moi*, Spanish *disculpe*). Some expressions may be tracked back to verbal multi-word expressions (English *please* < *if it please* (subjunctive), French *s'il vous plaît*), or to simpler adverbial multi-word expressions (Italian *per favore*, Spanish *por favor*, see Ben Khelil *et al.* 2022). Even at a semantic level, there are similarities in their origins. When addressing a second person, it is common to use verbs of excusing,

1 The transcriptions follow the conventions typically used for classical Greek and do not reflect the phonetic articulation of each word in the time period in which each text was written.

©2026 Lucía Madrigal Acero, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.06

i.e. verbs that are aimed at freeing the speaker from the guilt of disturbing the addressee (English *excuse me*, French *excusez-moi*, Spanish *disculpe*, Italian *scusi*, German *entschuldigen Sie*). Another common source are nouns meaning *favour* (Italian *per favore*, Spanish *por favor*, Arabic من فضلك *min faḍlak*), which may also appear in different formulae, such as SMG κάνε μου την χάρη *káne mou tin khári* "do me a favour." The phrase that I look into in this chapter is also semantically related to the notion of favour (ἀγάπη *agápē*), but differs from these examples, in that it has its origin in a support-verb construction (see Section 2).[2]

These strategies and phrases are typically used in oral exchanges, and it is difficult to guess what communication strategies were used in everyday Greek from the literature that has been preserved. For instance, we are uncertain of the formulae used to say *please* (but see La Roi 2025) or *thank you* in spoken language (Denizot 2022). At best, we can guess through parallels with SMG or the formulaic forms of address found in letters. However, it is unlikely that we date these utterings correctly, given the complexity and the turnabouts in the historical development of SMG, and the archaising and literary character of forms of address in ancient letters.

The aim of this chapter is to explore the development of the phrase ποιέω ἀγάπην *poiéō agápēn* into a marker of politeness in post-classical Greek. Other research has explored the development of specific verbs or verbal phrases in Greek into pragmatic markers (La Roi 2022, 2025; Fendel 2024). This chapter is, to my knowledge, the first inquiry into the use of a support-verb construction, ποίησον ἀγάπην *poíēson agápēn*, as a politeness marker that reduces the illocutive force of the following command or question. The phrase ποίησον ἀγάπην *poíēson agápēn* may have been overlooked in literary texts as a pragmatic marker due to its short life and its similarity to the realisative construction ποιέω ἀγάπην *poiéō agápēn* "to do a favour." However, previous work on support-verb constructions in Greek papyri already noted its prevalence in non-literary texts, which may push back the date for the beginning of the pragmaticalisation of ποίησον ἀγάπην *poíēson agápēn*. For instance Fendel (2022: 142) provides a reference for the phrase ποιέω ἀγάπην *poiéō agápēn* in a fourth-century Christian papyrus. The Christian context of this finding is relevant in view of the kinds of texts in which our literary examples are found, which are also Christian (see Section 2).

This chapter intends to explore three research questions:

1. What is the meaning of ποιέω ἀγάπην *poiéō agápēn*?
2. What are the indicators of the pragmaticalisation of ποίησον ἀγάπην *poíēson agápēn*?

2 Research on pre-modern Greek pragmatic markers is an ongoing process. Before ποίησον ἀγαπήν *poíēson agápēn*, which is a late formula, there were different means to ask for a favour or for information, such as parenthetical ἱκετεύω *hiketeúō* "I beg." This and other formulae for "please" in archaic and classical Greek have been discussed by La Roi (2025).

3. To what extent is ποίησον ἀγάπην *poíēson agápēn* a pragmaticalised phrase, adopting a gradient rather than categorical view of pragmaticalisation?

In the following sections, I describe my working data (Section 2) and the changes undergone by ποιέω ἀγάπην *poiéō agápēn* in its progressive acquisition of pragmatic functions. Firstly, I describe the semantic changes (Section 3). Secondly, I provide a syntactic analysis (Section 4), focussing on the (absence of) ossification of morphological number in the expression ποίησον ἀγάπην *poiēson agápēn* (Section 4.1), the syntactic position in which it tends to appear (Section 4.2), its degree of syntactic integration in the clause (Section 4.3), and its main pragmatic functions (Section 4.4). Lastly, I provide a synthesis of the main points of this chapter and make some concluding remarks (Section 5).

2. Research data

The data that I have analysed are the results of searching for ποιέω *poiéō* in co-occurrence with ἀγάπην *agápēn* in the *Thesaurus Linguae Graecae* (https://stephanus.tlg.uci.edu/).[3] This allows to track the semantic and syntactic development of the different meanings and contexts of use of the polysemic multi-word expression ποιέω ἀγάπην *poiéō agápēn*: "to love; to perform an act of love; to make love; to make peace; to celebrate a banquet; to do a favour" or "please" (see Section 3). Originally, ποιέομαι ἀγάπην *poiéomai agápēn* was used as a support-verb construction. Support-verb constructions are semi-compositional verb-noun phrases in which the noun has predicative properties, that is, it governs its own arguments, and the verb performs the function of an operator, which provides the nouns' arguments with the syntactic slots that they need for their realisation in a sentence. This makes support-verb constructions a peculiar kind of multi-word expressions as defined in the PARSEME Universal Annotation Guidelines for Multi-Word Expressions.[4] Multi-word expressions are non-compositional combinations that include at least two different lexemes. However, non-compositionality is a scalar property and, because support-verb constructions are only semi-compositional, they behave somewhat like prototypical multi-word expressions, but also like free constructions in certain aspects. A strict definition of prototypical support-verb constructions, which is the approach that I take in this chapter, includes only collocations formed by a semantically empty and syntactically transitive verb and its object (*to throw a tantrum, to give a hug, to make a request*) (Jiménez López 2016; Mel'čuk 2023;

3 The dataset is accessible here: https://doi.org/10.5281/zenodo.17491996
4 See https://parsemefr.lis-lab.fr/parseme-st-guidelines/2.0/

Fendel 2025).[5] One of the features that distinguishes support-verb constructions from other phraseological units is that they show, to a certain extent, some syntactic flexibility (Gross 2004; Fendel 2023). For example:

(1) I **don't** have a clue.

(2) I have **no** clue.

(3) Jane did a **brilliant** job.

(4) Jane did her job **brilliantly**.

(5) Mary **took a walk** down the riverside.

(6) The **walk that** Mary **took** down the riverside was very pleasant.

Examples (1) and (2) evidence the mobility of negation in support-verb constructions (Fendel 2023; 2025). Since the verb and the noun form a semi-lexical unit, the whole phrase is negated whether negation lies within the scope of the verb, (1), or the noun, (2). A similar phenomenon affects the complementation of support-verb constructions, (3) and (4). The same complement can be formulated as accompanying the verb, i.e. as an adverb, (4), or the noun, i.e. as an adjective, (3). Since the support-verb construction functions as a lexical unit, the modification of the verb or the noun affects the whole phrase, rather than just one of its component lexemes. Other evidence of this syntactic flexibility includes the possibility of extracting the noun from the phrase and replacing it with a relative pronoun, (6).

When syntactic flexibility is lost, one must consider whether the support-verb construction has experienced some degree of ossification—hence moving into the category of idioms—or whether another factor is driving changes in the phrase. This decrease in syntactic flexibility indicates that, at least from the fourth century CE onwards, ποιέω ἀγάπην *poiéō agápēn* no longer behaves as a regular support-verb construction, but rather as a pragmatic marker (Section 3).

The current study is corpus-based and focusses on the analysis of Greek texts extracted from the *Thesaurus Linguae Graecae*. The selection process has been manual. I have carried out a proximity search in the whole corpus of the *Thesaurus Linguae Graecae* in order to find all the instances of ποιέω *poiéō* and ἀγάπην *agápēn* within a seven-word range, discarding false positives which were generally cases in which ποιέω *poiéō* and ἀγάπην *agápēn* appear in separate clauses or cases in which ἀγάπη *agápē* is used in a case other than the accusative. The search intended to also look into other uses of ποιέω ἀγάπην *poiéō agápēn* in different contexts. The findings allow for the identification of at

5 There is a broader definition that includes collocations with semantically full verbs and different syntactic realizations (*to engage in conversation*, *to put something to an end*, *to keep a promise*) (Vivès 1984; Baños 2014; Jiménez López 2021). On these concepts, see also Alonso Ramos (2004); Mel'čuk (2023).

least eight uses of ποιέω ἀγάπην *poiéō agápēn* (see Section 3).

The first time that pragmaticalised ποίησον ἀγάπην *poíēson agápēn* appears in Greek literary texts—i.e. in the *Thesaurus Linguae Graecae*—is in the fourth century CE. After that point, its use becomes widespread until the sixth century CE, when its decline begins. However, ποίησον ἀγάπην *poíēson agápēn* remains present in the literature in later centuries, probably because of the conservative character of written Greek and its liking of the imitation of previous literary models. The full list of works which contain the (partially) pragmaticalised phrase ποίησον ἀγάπην *poíēson agápēn*, together with their analysis, is included in the dataset (Madrigal Acero 2025). Here, I provide an abbreviated overview, Table 6.1, with all the instances of pragmaticalised ποίησον ἀγάπην *poíēson agápēn* or ποιήσετε ἀγάπην *poiḗsete agápēn* that I have identified (see Sections 3.7 and 3.8), together with their morphological, syntactic and pragmatic analysis. I have left out of Table 6.1 the contexts in which ποιέω ἀγάπην *poiéō agápēn* does not show traces of pragmaticalisation, hence why Table 6.1 begins with the examples from the fourth century CE. As I will argue below, the recruitment of pragmatic properties only occurs in the phrase with the verb in the imperative aorist, which is evidence of the ossification of a construction which was previously more flexible from both the syntactic and morphological perspectives. This feature also renders more systematic the distinction between collocations (false positives) and pragmatic markers (real positives).

Date	**Adressee**		**Position**		**Syntactic integration**			**Function**	
	Singular	Plural	1st	2nd/ other	Main verb	καί *kaí*	Juxtaposition	Request	Question
4th c. CE	33	7	35	5	1	18	21	37	3
5th c. CE	18	1	17	2	1	11	7	19	0
6th c. CE	34	4	28	10	3	15	20	27	11
7th c. CE	7	2	8	1	0	5	4	7	2
8th c. CE	1	0	1	0	0	1	0	0	1
10th c. CE	4	0	2	2	0	2	2	2	2
13th c. CE	10	0	4	6	1	9	0	9	1
18th c. CE	1	0	0	1	0	1	0	1	0
Total	108	14	95	27	6	62	54	102	20

Table 6.1 Instances of pragmaticalised ποίησον ἀγάπην/ποιήσετε ἀγάπην *poíēson agápēn*/ *poiḗsete agápēn* in the *Thesaurus Linguae Graecae* in chronological order.

Most of the instances of ποίησον ἀγάπην *poíēson agápēn* appear in texts written between the fourth and the sixth centuries CE. Some of the later uses appear in texts that rework or even quote the information from earlier sources. These

later uses are to be considered remains of a receding phrase that peaked in the literature of the period from the fourth to sixth centuries CE, and which was already in use in documentary papyri (see Fendel 2022: 142). However, I have not found examples of ποίησον ἀγάπην *poíēson agápēn* in documentary papyri dated prior to the fourth century CE. I have not excluded the more recent examples (from the seventh century CE onwards, see Table 6.1) from my analysis because not all of them are secondary in this sense, and it is important to see the whole picture, where older phrases keep being brought into later texts or speech due to the use of earlier sources or due to the Greek diglossia (Toufexis 2008; Horrocks 2014). Furthermore, the decline of this phrase may be linked to the prevalence of different politeness strategies in post-classical Greek. For instance, in Greek papyri collocations with the noun χάρις *kʰáris* 'grace' are very productive (Fendel 2022). Nevertheless, the question whether the gradual disappearance of ποίησον ἀγάπην *poíēson agápēn* is due to lexical substitution with another phrase or with other expressions of politeness is beyond the scope of this chapter.

Apart from chronology, another factor has proved relevant in the distribution of ποίησον ἀγάπην *poíēson agápēn*: sociocultural context. The texts in which it is found are Christian—with very few exceptions, such as the *Erotopaignia*—which aligns with Goutsos' (2017) findings that sociolinguistics impacts greatly the distribution of certain pragmatic markers. Most of the examples are found in hagiographical or ecclesiastical texts. Even the documentary papyri that include this phrase belong to a Christian context. The example cited in the introduction (Fendel 2022: 142) comes from the archive of Apa Paeious, who was the head of a fourth-century Meletian community.

The table also presents the variables that I have included in my analysis of ποίησον ἀγάπην *poíēson agápēn*: addressee(s), position, degree of syntactic integration and pragmatic function. All of them are discussed in depth in Section 4.

The next sections develop the semantic and syntactic changes that ποιέω ἀγάπην *poiéō agápēn* underwent on its path to pragmaticalisation. The way I understand pragmaticalisation in this chapter is as a process involved in language change which shares a lot of properties with grammaticalisation and is, in fact, part of the process of grammaticalisation. It is not the purpose of this chapter to discuss the distinction between grammaticalisation and pragmaticalisation, for which I refer to the chapters that I have used (Degand and Evers-Vermeul 2015; La Roi 2022). More importantly, Degand and Evers-Vermeul (2015: 68) sum up the semantic and syntactic processes undergone by pragmaticalising units. The semantic changes include:

> (i) the evolution from referential to discourse-orientated and/or interpersonal meaning including semantic bleaching through pragmatic strengthening, (ii) persistence of the original lexico-grammatical meaning, (iii) subjectification, and to a certain extent (iv) intersubjectification (Degand and Evers-Vermeul 2015: 68–69).

On the syntactic level, they list "(i) scope extension, (ii) decategorialization, and (iii) syntactic optionality" (Degand and Evers-Vermeul 2015: 69). These are some of the characteristics of pragmaticalisation that affect the change in ποίησον ἀγάπην *poíēson agápēn*.

The next section starts with a description of the semantic uses of ποιέω ἀγάπην *poiéō agápēn* that I have identified within the corpus and ends with an introduction to ποίησον ἀγάπην *poíēson agápēn* as a pragmatic marker.

3. Semantic changes

The phrase ποιέω ἀγάπην *poiéō agápēn* can be used in several different ways, which are, on the one hand, due to the polysemy of the noun ἀγάπη *agápē* and, on the other, to its progressive recruitment of pragmatic properties. In this section, I introduce the eight main meanings of ποιέω ἀγάπην *poiéō agápēn* present in the *Thesaurus Linguae Graecae* and trace the semantic relationship between them. There are three major sets in which all these meanings are articulated. The first one includes the only support-verb construction in the group (Section 3.1). The second set includes the instances in which ποιέω ἀγάπην *poiéō agápēn* is used as a realisative construction with different meanings (Sections 3.2 to 3.6). The last set is formed by the pragmaticalised uses of ποίησον ἀγάπην *poíēson agápēn* (Sections 3.7 and 3.8).

3.1 ποιέομαι ἀγάπην *poiéomai agápēn* "to love" (≈ ἀγαπάω *agapáō)*

The verb ἀγάπάω *agapáō* "to love" is morphologically derived from the noun ἀγάπη *agápē*, whose core meaning is "love" (LSJ, DGE, Bailly, s.v.). Therefore, the first meaning of the phrase ποιέομαι ἀγάπην *poiéomai agápēn* is "to love." Example (7) illustrates this by creating a parallel sentence with the synthetic expression in the first clause and the analytic form in the second one.

(7) Basilius Ancyranus, *De virginitate—Patrologia Graecae* 30.744[6]

ἀγαπᾷν οὖν δεῖ τοὺς πρὸς τὸν νυμφίον παρακαλοῦντας, ἀλλὰ δεῖ τὴν **ἀγάπην** δικαίαν καὶ μὴ συγκεχυμένην **ποιεῖσθαι**

agapâin *oûn deî toùs pròs tòn nymp*^{h}*íon parakaloûntas, allà deî tèn* ***agápēn*** *dikaían kaì mè synkek*^{h}*yménēn* ***poieîst***^{h}***ai***

"It is necessary to **love** those who exhort the groom, but it is necessary to **love** justly and not unfairly."

6 Where there is a known author, their name is given first in Latin, according to the *Thesaurus Linguae Graecae* (https://stephanus.tlg.uci.edu/), followed by the work title. The titles of the works included here are those given on *Thesaurus Linguae Graecae*.

Example (7) illustrates a context in which the synthetic verb ἀγαπᾶν *agapân* and the analytic expression ποιεῖσθαι ἀγάπην *poieîst^{h}ai agápēn* are used as synonyms. The phrase ποιέομαι ἀγάπην *poiéomai agápēn* is here a support-verb construction (Alonso Ramos 2004; Jiménez López 2016; Baños et al. 2022; Mel'čuk 2023; Fendel 2025). Some differences between the support-verb construction and the simplex verb include the greater syntactic flexibility of the support-verb construction and its use of the middle voice. Firstly, the noun in the support-verb construction in (7) is determined by the article την *tèn* and modified by the adjective δικαίαν *dikaían* and the participle μὴ συνκεχυμένην *mè synkek^{h}yménēn*. None of this is possible when the speaker chooses the synthetic expression. If anything, it is possible to modify the simplex verb with an adverb, but it seems difficult to adverbialise a complex expression such as a participial clause. Secondly, the support verb ποιέομαι *poiéomai* needs to be used in the middle voice due to the co-referentiality of the first argument of the noun ἀγάπην *agápēn* and the syntactic subject of the verb ποιεῖσθαι *poieîst^{h}ai* (Cock 1981; Jiménez López 2016; Fendel 2025). In this specific case, the clause is in fact impersonal, which gives the statement general validity.

3.2 ποιέω ἀγάπην *poiéō agápēn* "to perform an act of love" (realisative construction)

(8) Barsanuphius et Joannes, *Quaestiones et responsiones ad coenobitas*, 503.25

καὶ **ποίησον τὴν ἀγάπην** λαλῶν τῷ πλησίον, διάκρινον εἰ μὲν μετὰ ταπεινώσεως καὶ διακρίσεως ἀταράχως λαλεῖς, λάλησον, εἰ δέ ἐστι τίποτε μὴ τοιοῦτον, ἄγξον τὸν λογισμὸν καὶ καταίσχυνον αὐτὸν καὶ παύεται

*kaì **poíēson tèn agápēn** lalôn tôi plēsíon, diákrinon ei mèn metà tapeinṓseōs kaì diakríseōs atarák^{h}ōs laleîs, lálēson, ei dé esti típote mè toioûton, ánkson tòn logismòn kaì kataísk^{h}ynon autòn kaì paúetai*

"And **practice your love** by speaking to your neighbour, discern; if you speak with humility and discernment, speak, but if there is something else unlike this, silence your reason, dishonour it and cease."

Still related to the core meaning of ἀγάπη *agápē*, there is a different phrase: ποιέω ἀγάπην *poiéō agápēn* "to practice love or charity". Morphologically, this phrase differs from the support-verb construction in that it is used in the active voice. This is an indication that the phrase is not used as a support-verb construction, but rather as something else: a realisative construction. The definition for realisative constructions varies from author to author and we are still far from a consensus on the matter (Alonso Ramos 2004; Mel'čuk 2023; Fendel 2025). The definition I follow in this chapter is that of Alonso Ramos (2004: 113–114), which focusses on the semantics of the verb. Whereas a support verb is semantically empty, meaning that it adds no meaning to the construction

and performs the function of a syntactic operator, in a realisative construction the verb retains a part of its original semantics, and conveys the notion that the purpose of the existence of the noun has been fulfilled. A realisative verb can be glossed as to-carry-out-the-purpose-of-the-noun. There are different ways to understand this definition, which becomes clear through the examples that each author gives. For instance, Alonso Ramos (2004: 114) gives the examples *to fulfill/keep a promise*. The goal of a *promise* is not to be broken, but rather to be *fulfilled* or *kept*. Hence, there is an obvious semantic difference between the prototypical support-verb construction *to make a promise*, where the verb is semantically empty, and the realisative-verb constructions *to fulfill*/*keep a promise*, where it is not.[7] In the case of ποιέω ἀγάπην *poiéō agápēn*, ποιέω *poiéō* retains its meaning of "making, manufacturing", as in "creating a physical reality that materialises the abstract concept of ἀγάπη *agape*" and implies that the objective of said ἀγάπη *agápē* is fulfilled, that is, helping another person such as to make them feel taken care of and loved. Therefore, there is a syntactic explanation for the verb being used in the active: the subject of ποιέω *poiéō* is not an Experiencer of ἀγάπη *agápē*—i.e. the first argument of the noun—but rather an Agent that carries out an action. So, while there are other models that define support-verb constructions and realisative constructions differently, how I use the distinction between realisative and support verbs is especially relevant in Greek because it is supported by the voice alternation. While support-verb constructions use *poiéomai* in the middle due to the co-referentiality between the first argument of the predicative noun and the subject of the support verb, realisative constructions use ποιέω *poiéō* because the active voice implies that the subject is more Agent-like and not necessarily coreferential, and that the verb is semantically closer to the core meaning of ποιέω *poiéō* "to create" than to the semantically empty support verb ποιέομαι *poiéomai* "to do".[8]

I would also like to underscore another methodological caveat, which is that, in

7 Different views have been taken on the issue. While Mel'čuk (2023: 225) broadly follows Alonso Ramos' (2004) definition, he expands it to expressions with what he calls quasi-predicative nouns. Fendel (2025) takes a different stance on the issue by following Mel'čuk's (2004) distinction of two main groups of verbs, support verbs and verbs of realisation, which she then applies to classical Greek as a descriptive category. Her definition of verbs of realisation also includes what some researchers have called *verbes appropriés* (Gross 1999) and what some researchers have called aspectual support-verb extensions and causative/diathetic support-verb extensions (Alonso Ramos 2004; Baños 2014; Jiménez López 2021; Mel'čuk 2023). In this chapter I adhere to Alonso Ramos' (2004) definitions.

8 On top of this, the middle voice is a receding category in post-classical Greek, and in SMG it only survives in *media tantum* verbs. Holton et al. (2019: 1434) state: "In the period covered by this Grammar [sc. from the eleventh to the eighteenth century], the middle voice, which had begun to be replaced by the passive in the Koine […], had long disappeared and only occurs in residual forms in texts belonging to mixed or higher stylistic registers." However, in another section, they date the beginning of this process to an earlier period: "In the Koine the middle voice had started to fall into disuse, and to be replaced with passive or active morphology" (Holton et al. 2019: 1375).

this chapter, I do not apply the labels realisative or support (verb) to specific verbs, as if they were properties inherent to them, but to constructions. This implies that a single verb may function as a support verb in a certain context, but as a different kind of verb in another. In other words, I believe that the same verb has a set of functions, of which one is activated by the specific context in which it is used.

3.3 ποιέω ἀγάπην *poiéō agápēn* "to make love" (realisative construction)

(9) Epiphanius, *Panarion*, 1.280.23

ἐκ τούτου δὲ συμποσιάσαντες καὶ ὡς ἔπος εἰπεῖν τὰς φλέβας τοῦ κόρου ἐμπλήσαντες ἑαυτῶν εἰς οἶστρον τρέπονται. καὶ ὁ μὲν ἀνὴρ τῆς γυναικὸς ὑποχωρήσας φάσκει λέγων τῇ ἑαυτοῦ γυναικὶ ὅτι «ἀνάστα, **ποίησον τὴν ἀγάπην** μετὰ τοῦ ἀδελφοῦ». οἱ δὲ τάλανες μιγέντες ἀλλήλοις καί, ὡς ἐπὶ ἀληθείας αἰσχύνομαι εἰπεῖν τὰ παρὰ τοῖς αὐτοῖς αἰσχρὰ πραττόμενα [...], ὅμως οὐκ αἰσχυνθῶ λέγειν ἃ αὐτοὶ ποιεῖν οὐκ αἰσχύνονται, ἵνα κατὰ πάντα τρόπον φρῖξιν ἐργάσωμαι τοῖς ἀκούουσι τὰ παρ' αὐτῶν τολμώμενα αἰσχρουργήματα, ...

*ek toútou dè symposiásantes kaì ʰōs épos eipeîn tàs pʰlébas toû kórou emplḗsantes ʰeautôn eis oîstron trépontai. kaì ʰo mèn anḕr tês gynaikòs ʰypokʰōrḗsas pʰáskei légōn têi ʰeautoû gynaikì ʰóti «anásta, **poiḗson tḕn agápēn** metà toû adelpʰoû». ʰoi dè tálanes migéntes allḗlois kaí, ʰōs epì alḗtʰeias aiskʰýnomai eipeîn tà parà toîs autoîs aiskʰrà prattómena [...], ʰómōs ouk aiskʰyntʰṓ légein ʰà autoì poieîn ouk aiskʰýnontai, ʰína katà pánta trópon pʰrîksin ergásōmai toîs akoúousi tà par' autôn tolmômena aiskʰrourgḗmata, ...*

"After this, having feasted and, so to say, having filled themselves with excess, they turn to frenzy. And the man, withdrawing from his wife, speaks telling his own wife that 'Get up, **make love** with the brother.' The wretches, mingling with one another and—I am ashamed of saying the evil deeds done by them,—[...] I will not be ashamed of saying what they are not ashamed to do, in order to frighten those who listen to the shameful acts that they dared to perform..."

(10) *Erotopaignia (e manuscripto Londinense)*, 3.315

ἑκατὸν χρόνοι καὶ ἂν διαβοῦν, φιλὶν μὲ θέλεις δώσει, καὶ τί με τυραννεῖς, κυρά, μὲ τὴν γλυκειά σου ἀγάπην, μὲ τὸ αὔριον. μὲ τὸ σήμερον, μὲ τὸ καρτέρει, αὐθέντη; ἐπλήρωσεν τὸ στοίχημα· ἔλα, κυρά, ἂς φιλοῦμεν, ἔλα, κυρά, ἂς **ποιήσωμεν** πολλῶν χρονῶν **ἀγάπην**

*ʰekatòn kʰrónoi kaì àn diaboûn, pʰilìn mè tʰéleis dṓsei, kaì tí me tyranneîs, kyrá, mè tḕ glykeiá sou agápēn, mè tò aúrion. mè tò sḗmeron, mè tò kartérei, autʰéntē; eplḗrōsen tò stoíkʰēma· éla, kyrá, às pʰiloûmen, éla, kyrá, às **poiḗsōmen** pollôn kʰrónōn **agápēn***

"Even if a hundred years should pass, you want to give me a kiss. Why do you torment me, my lady, with your sweet love, with tomorrow, with today, with patience, sovereign? The price has been paid. Come on, my lady, let's kiss. Come on, my lady, let's **make love** for many years."

Also related to the core meaning of ἀγάπη *agápē* "love" is another sense of ποιέω ἀγάπην *poiéō agápēn* "to make love." Morphologically, semantically and syntactically, this phrase behaves almost like (ii), but refers exclusively to having intercourse, which may be considered a specific "act of love". The difference between (ii) and (iii) is clear from the contexts in which they are used. (ii) is used to refer to the Christian notion of "charity" or "love towards other human beings," whereas (iii) is used in a sexual sense. Example (9) condemns the action of ποιέω ἀγάπην *poiéō agápēn* as being a shameful act, whereas (10) shows its use in erotic poetry.

3.4 ποιέω ἀγάπην *poiéō agápēn* "to make peace, to make an alliance" (realisative construction)

Another construction is ποιέω ἀγάπην *poiéō agápēn* "to make peace". This construction also takes the active voice, which is a clue that it is in fact a realisative construction. Example (11) addresses politics, so a political interpretation of the term ἀγάπη *agápē* is required.

(11) *Collatio quinta. Novellae constitutiones annorum* 1204–1453, 6.28

ἵνα μετὰ τὸ συνδεθῆναι αὐτοὺς εἰς ἀγάπην μετὰ τῆς βασιλείας μου οὐ μὴν **ποιήσωσιν** οὗτοι μετὰ ἄλλου γένους **ἀγάπην** κατὰ τῆς βασιλείας μου

*ʰína metà tò syndetʰênai autoùs eis agápēn metà tês basileías mou ou mèn **poiḗsōsin** ʰoûtoi metà állou génous **agápēn** katà tês basileías mou*

"So that, after having begged for peace with my kingdom, they will not **make an alliance** with another people against my kingdom."

Kingdoms do not have the ability to experience love, so the closest thing to love between two kingdoms is friendship or an alliance between them, analogous to two people who show ἀγάπη *agápē* to one another. The first argument of ἀγάπη *agápē* "alliance" is the entity who makes or holds the alliance, in this case, the people referred to by αὐτούς *autoús* "them," which is the logical subject of ποιήσωσιν *poiḗsōsin*. In other words, the first argument of ἀγάπη *agápē* and the subject of ποιέω *poiéō* are coreferential, yet the verb is used in the active because this is not the support-verb construction ποιέομαι ἀγάπην *poiéomai agápēn* "to love", which refers to an emotion. Here, the subject has more agentive characteristics: rather than experiencing something, the subject is *making* something, i.e. a new political situation which requires negotiation and active involvement.

It seems, however, that there is a historical instability in other similar phrases. For instance, ποιέομαι συμμαχίαν *poiéomai symmakʰían* and ποιέομαι φιλίαν *poiéomai pʰilían* "to make an alliance" were widely used collocations in classical Greek.[9] These phrases co-existed with the causative phrase συμμαχίαν ποιέω

9 The main dictionaries provide some references for these support-verb

symmakʰían poiéō and φιλίαν ποιέω *pʰilían poiéō* "to cause the establishment of an alliance", which was morphologically marked using the active voice. However, in certain contexts, the distinction between making an alliance and causing an alliance is unclear, (12) and (13).[10] This in turn might have led to a neutralisation of the voice alternation specifically with these nouns, which often made the distinction between support-verb construction, realisative construction, and causative construction blurry due to their semantics. Furthermore, the confusion between active and middle may be considered an indicator that these constructions were in fact neither support-verb constructions nor causative constructions, but something in between: realisative constructions.

(12) Isocrates, *Speech*, 5.44.4

ἡμεῖς δὲ καὶ πρὸς ἐκείνους μόνοι τῶν Ἑλλήνων **ποιησάμενοι συμμαχίαν** συναίτιοι τῆς σωτηρίας

*ʰēmeîs dè kaì pròs ekeínous mónoi tôn ʰEllḗnōn **poiēsámenoi symmakʰían** synaítioi tês sōtērías*

"We, who were the only Greeks that **made an alliance** with them, were the cause of their salvation."

(13) Diodorus Siculus, *Bibliotheca Historica*, 12.5.2.3

ἐπὶ δὲ τούτων Μεγαρεῖς μὲν ἀπέστησαν ἀπὸ Ἀθηναίων, καὶ πρὸς Λακεδαιμονίους διαπρεσβευσάμενοι **συμμαχίαν ἐποίησαν**

*epì dè toûtōn Megareîs mèn apéstēsan apò Atʰēnaíōn, kaì pròs Lakedaimoníous diapresbeusámenoi **symmakʰían epoíēsan***

"Then, the Megareans removed from the Athenians, and, after sending ambassadors to the Lacedaemonians, they **made an alliance**."

Between ποιέομαι συμμαχίαν *poiéomai symmakʰían*, (12), and ποιέω συμμαχίαν *poiéō symmakʰían*, (13), there seem to be no semantic or syntactic differences. In (12), the subject of ποιησάμενοι *poiēsámenoi* and the Agent of συμμαχίαν *symmakʰían* is ἡμεῖς *ʰēmeîs*, that is, there is co-referentiality, which justifies the middle voice. However, (13) seems to behave equally, but opts for the active. From a semantic perspective, the distinction between a neutral "to make an alliance" and the explicitly causative "to cause an alliance" seems far from clear-cut, which accounts for the neutralisation of the voice alternation in phrases meaning "to make an alliance".

constructions. For ποιέομαι συμμαχίαν *poiéomai symmakhían*, see Herodotus, *Histories*, 5.73, Xenophon, *Hellenica*, 3.2.21, *Inscriptiones Graecae* 2(2).43.26 (LSJ, s.v.). For ποιέομαι φιλίαν *poiéomai philían*, see Xenophon, *Memorabilia*, 2.6.29 (LSJ, s.v.).

10 For examples from the *Dictionary of Ancient Greek Collocations* (DiCoGra): https://dicogra.iatext.ulpgc.es/

Another piece of evidence for the use of ποιέω ἀγάπην *poiéō agápēn* "to make an alliance" as a realisative construction is (14):

(14) Vitae et miracula Nicolai Myrensis, *Epitome (e cod. Paris. gr. 1034)*, 8.3

ἰδόντες δὲ οἱ τρεῖς ἐκεῖνοι ἄρχοντες τὸ γεγονὸς ἐθαύμασαν. καὶ εὐλογηθέντες ὑπ' αὐτοῦ, πρὸς τὴν Φρυγίαν ἀφίκοντο. καὶ πορευθέντες ἐκεῖσε, **ἀγάπην καὶ εἰρήνην ἐποίησαν**

idóntes dè ʰoi treîs ekeînoi árkʰontes tò gegonòs etʰaúmasan. kaì eulogētʰéntes ʰyp' autoû, pròs tèn Pʰrygían apʰíkonto. kaì poreutʰéntes ekeîse, ***agápēn kaì eirḗnēn epoíēsan***

"Upon seeing it, those three rulers marvelled at what had happened. And having been praised by him, they went to Phrygia. And having arrived there, they **made an alliance and peace**."

In this text, three rulers witness a miracle and put an end to the conflict among their territories. The noun ἀγάπην *agápēn* is coordinated with εἰρήνη *eirḗnē*, which is a partial synonym in this context. Already in classical Greek, εἰρήνη *eirḗnē* could be used both with ποιέομαι *poiéomai* and with ποιέω *poiéō*. However, from a semantic perspective, the distinction between the neutral support-verb construction and the causative construction is not always obvious. In (14), εἰρήνη *eirḗnē* and ἀγάπη *agápē* are combined with the active ποιέω *poiéō*, which points at them being in fact realisative constructions, rather than support-verb constructions.

3.5 ποιέω ἀγάπην/ἀγάπας *poiéō agápēn/agápas* "to celebrate a banquet" (realisative construction)

(15) *Apophthegmata patrum (collectio alphabetica)*, 328.15

ἀδελφὸς ἦλθε πρὸς τὸν ἀββᾶν Ποιμένα, καὶ λέγει αὐτῷ· Σπείρω τὸν ἀγρόν μου, καὶ ποιῶ ἐξ αὐτοῦ ἀγάπην

adelpʰòs êltʰe pròs tòn abbān Poiména, kaì légei autôi· speírō tòn agrón mou, kaì poiô eks autoû agápē

"A brother went to abbot Poimen and told him: 'I sow my field and make a banquet out of it.'"

(16) Joannes III, *Synagoga L titulorum*, 3.146

Ὅτι οὐ δεῖ ἐν τοῖς κυριακοῖς ἢ ἐν ταῖς ἐκκλησίαις **τὰς** λεγομένας **ἀγάπας ποιεῖν**, καὶ ἔνδον ἐν τῷ οἴκῳ ἐσθίειν καὶ ἀκούβιτα στρωννύειν

ʰóti ou deî en toîs kyriakoîs è en taîs ekklēsíais tàs legoménas ***agápas poieîn****, kaì éndon en tôi oíkōi estʰíein kaì akoúbita strōnnýein*

[On the prohibition of interacting with the excommunicated] "That one must not **celebrate the** so-called **banquets** on Sundays or in the churches, or eat inside the house and lay beds."

(17) Matthaeus Blastares, *Collectio alphabetica*, pin.Alpha

πῶς δεῖ **τὰς** λεγομένας **ἀγάπας ποιεῖν**, ἤτοι συμπόσια

*pôs deî tàs legoménas **agápas poieîn**, étoi sympósia*

"How the so-called **banquets**, or symposia, must be **celebrated**."

In Christian Greek, a new meaning of ἀγάπη *agápē* emerged which can also be traced back to "love." Thayer (1887, s.v.) defines ἀγάπαι *agápai*, in the plural, as

> love-feasts, feasts expressing and fostering mutual love which used to be held by Christians before the celebration of the Lord's supper, and at which the poorer Christians mingled with the wealthier and partook in common with the rest of food provided at the expense of the wealthy.

So, in certain cases, ἀγάπη *agápē* is used to refer to a ritual which was used for the expression of mutual affection and solidarity in a religious context. As in (ii) to (iv), the noun refers to a specific realisation of ἀγάπη *agápē*, and the verb is used in the active to indicate that the subject is not an Experiencer of ἀγάπη *agápē*, but rather a more prototypical Agent. Furthermore, this meaning is more typically used in the plural, (16) and (17), which also leads to interpreting the predicate as an iterative activity, i.e. a ritual that is repeated consistently over time. However, when the speaker refers to a specific ἀγάπη *agápē*, rather than the performance of ἀγάπαι *agápai* in general, the noun may also be used in the singular, (15). Typically, the context provides the necessary information to identify this value of ποιέω ἀγάπην *poiéō agápēn*. In (15), the text speaks about how a monk takes care of a garden where he grows the food with which he is then able to provide for his guests at an ἀγάπη *agápē*. In (16), there are specific signals that the text talks about the religious ritual, such as the specific mention of ταῖς ἐκκλησίαις *taîs ekklēsíais* "the churches" or the verb ἐσθίειν *est^{h}íein* "to eat." Example (17) provides a synonym for *ἀγάπαι agápai*, which is συμπόσια *sympósia*.

3.6 ποιέω ἀγάπην *poiéō agápēn* "to do a favour" (realisative construction)

(18) *Assisiae (e cod. Par. Suppl. gr. 465)*, 81.5

ἐὰν λάχῃ ὅτι εἷς ἄνθρωπος θελήσει νὰ **ποίσῃ** (*sic*) χρηστότηταν καὶ **ἀγάπην** ἑνοῦ του φίλου, καὶ δανείζει του τὸ ἄλογόν του [...]

*eàn lák^{h}êi ^{h}óti ^{h}eîs ánt^{h}rōpos t^{h}elḗsei nà **poísēi** k^{h}rēstótētan kaì **agápēn** ^{h}enoû tou p^{h}ílou, kaì daneízei tou tò álogón tou [...]*

"If by chance a man wants to be kind and **do a favour** to a friend, and lends him his horse [...]."

(19) Daniel Scetiota, *De mendico caeco*, 8

ποίησον εἰς ἐμὲ ἀγάπην, ἀδελφέ, ὅτι οὐκ ἔχω πόθεν ἀγοράσαι μοι βαΐα τοῦ καμεῖν καὶ τραφῆναι

***poíēson** eis emè **agápēn**, adelp^{h}é, ^{h}óti ouk ék^{h}ō pót^{h}en agorásai moi baḯa toû kameîn kaì trap^{h}ênai*

"[A blind man begs to a passer-by] **Do me a favour**, brother, for I do not have the means to buy a rod to work and support myself."

Another meaning that can be derived from "perform an action to show appreciation or affection to someone" is "to do a favour". This sense of the phrase is common from the fourth century CE onwards, when it appears in literary texts for the first time.[11] It seems that the phrase was also common in the papyri of the same period (DGE, s.v.). In (18), ἀγάπην *agápēn* is coordinated with χρηστότηταν *k^{h}rēstótētan*, which is here used as a quasi-synonym. However, what tells that ἀγάπην *agápēn* here does not refer merely to any other kind of act of love, but to a specific favour, is the following sentence, where the author details that the favour consists of a man lending his horse to a friend. In (19), the favour that the speaker is asking for is for the addressee to buy him a rod.

In both examples, ποιέω ἀγάπην *poiéō agápēn* is used in the active voice. Early on, ποίησον ἀγάπην *poíēson agápēn* in the imperative will be pragmaticalised, but examples such as (18) and (19) display how it could also be used as the main verb of a sentence, (19), coordinated with another noun, (18), or complemented by a prepositional phrase, (19). This is an indication that the pragmaticalised phrase (ποίησον ἀγάπην *poíēson agápēn* "please") co-existed to a certain extent with the realisative construction (ποιέω ἀγάπην *poiéō agápēn* "to do a favour"). However, the (partially) pragmaticalised phrases are more frequent than the regular support-verb construction.

3.7 Pragmatic marker ("please" + request)

The sense of ποιέω ἀγάπην *poiéō agápēn* "to do a favour" is the source of the semi-pragmaticalised expression ποίησον ἀγάπην *poíēson agápēn* "do [me] a favour; please." There are certain indicators of an ongoing pragmaticalisation process, such as its acquisition of intersubjective meaning and, in most cases, its syntactic optionality. It is used in only two contexts: either to introduce a request (typically another imperative) or to introduce a question (typically a direct question, see Section 3.8). The syntactic changes undergone by ποίησον

11 The first appearance of the ποιέω ἀγάπην *poiéō agápēn* in this sense in the *Thesaurus Linguae Graecae* is Ephraem Syrus, *Sermones paraenetii ad monachos Aegypti*, 42.20, see (31). It is significant that this sense only appears in the aorist imperative, which hints at a process of ossification which, in turn, points at its acquisition of pragmatic meaning.

ἀγάπην *poíēson agápēn* are explored in depth in Section 4.

The main reason for separating uses (vi) and (vii) is that, despite semantics being persistent, there are differences in the syntax and the pragmatic use of the phrase. Whereas in (vi) the verb could appear in different parts of the sentence, in (vii) it can only appear in the left periphery introducing a demand in the imperative. While in (vi) it could take other complements, in (vii) it cannot. And, whereas in (vi) it could be conjugated in other moods and coordinated with other nouns, in (vii) it can only be used in the imperative with no noun co-ordination. The following examples evidence that, despite the semantic persistence, the phrase ποίησον ἀγάπην *poíēson agápēn* is not used freely within the clause anymore, but rather as a satellite, with an ossified morphological form and a fixed word order:

(20) Joannes Moschus, *Pratum spirituale*, 79.2929

τί κελεύεις, κύριε; Τότε λέγω αὐτῇ· **Ποίησον ἀγάπην**, ἐλθὲ ἕως ὧδε

*tí keleúeis, kýrie? tóte légō autêi· **poíēson agápēn**, elthè ʰéōs ʰôde*

"'What do you command, lord?' Then I tell her: '**Please**, come here.'"

(21) Pseudo-Gregentius, *Vita sancti Gregentii*, 5.165

διὰ τὸν κύριον **ποίησον ἀγάπην**, ἀδελφέ, καὶ θεράπευσον τοὺς ὀφθαλμοὺς τοῦδε τοῦ νεανία

*dià tòn kýrion **poíēson agápēn**, adelpʰé, kaì tʰerápeuson toùs opʰtʰalmoùs toûde toû neanía*

"For the Lord's sake, **do him a favour/please**, brother, and heal the eyes of this young man."

The main difference between (20) and (21) is that, whereas in (20) ποίησον ἀγάπην *poíēson agápēn* is syntactically disconnected from the main clause with ἐλθέ *eltʰé*, hence functioning as a parenthetical phrase, in (21) it is coordinated with the main verb, even if it is surrounded by other extra-clausal elements, such as διὰ τὸν Κύριον *dià tòn Kýrion*[12] "for the lord's sake" and ἀδελφέ *adelpʰé* "brother", so it could be argued that it is technically still a typical support-verb construction from a functional perspective. However, these are the two main options for the introduction of ποίησον ἀγάπην *poíēson agápēn* in a sentence: either coordinated by means of καί *kaí* "and" or simply juxtaposed to the main clause, which expresses a strong request, usually in the form of an imperative. It is very infrequent to find it used as the syntactic nucleus of the clause. It can therefore be concluded that it serves to mitigate the illocutionary force of the imperative, transforming it into a polite request.

12 The phrases διὰ τὸν Κύριον *dià tòn Kýrion* and διὰ τὸν θεόν *dià tòn tʰeón* appear frequently preceding or following ποίησον ἀγάπην *poíēson agápēn* as intensive markers.

3.8 Pragmatic marker ("please" + question)

In (vii), the meaning "to do a favour" is still clear, since the phrase ποίησον ἀγάπην *poíēson agápēn* is expanded by a request in the imperative which specifies the kind of favour that is being asked. Here, however, what follows ποίησον ἀγάπην *poíēson agápēn* is a question, whether direct, (22), or indirect, (23).

(22) Joannes Moschus, *Pratum spirituale*, 123.2985

ποίησον ἀγάπην, καλόγηρε, πόθεν με οἶδας;

***poíēson agápēn**, kalógēre, pót^{h}en me oîdas;*

"**Please**, Reverend, whence do you know me?"

(23) Theodorus Studites, *Laudatio S. Arsenii anachoretae*, 1.853.23

ποίησον ἀγάπην, ἀδελφέ, διὰ τὸν Κύριον, καὶ εἰπὲ ἡμῖν, ποῖον λογισμὸν ἔσχες ἐσθίων τὸν παξαμάν;

***poíēson agápēn**, adelp^{h}é, dià tòn Kýrion, kaì eipè ^{h}ēmîn, poîon logismòn ésk^{h}es est^{h}íōn tòn paksamán;*

"**Please**, brother, for the Lord's sake, tell us: 'What was your thought as you ate the biscuit?'"

The formulation of the clauses is very similar to the ones in (vii): ποίησον ἀγάπην *poíēson agápēn* in the left periphery, followed by a vocative and, optionally, the phrase διὰ τὸν Κύριον *dià tòn Kýrion* "for the lord's sake". In (22), the question is introduced directly after the vocative, but in (23) ποίησον ἀγάπην *poíēson agápēn* is coordinated with the imperative εἰπέ *eipé* "speak", which in turn introduces the question. In both examples, the aim of ποίησον ἀγάπην *poíēson agápēn* is to soften the illocutionary strength of the direct question, so as to introduce it more politely. In this context, the semantics of ἀγάπη *agápē* as "favour" is less clear: the speaker is not asking the addressee to carry out any specific action—if anything, the action of speaking in (23)—but rather to provide a certain piece of information. In (22), ποίησον ἀγάπην *poíēson agápēn* does not introduce any other verb, but rather just a direct question. This evidences that it is not at all integrated into the main clause but used as an attitudinal satellite. It indicates that the speaker does not wish to be overly invasive with their question, and that they will be thankful if the addressee decides to provide the requested information. In other words, this is the context in which pragmaticalisation, and more specifically intersubjectification, is the most evident. By intersubjectification, I understand the recruitment of pragmatic functions that focus on the reception of the message by the addressee, rather than the speaker. Intersubjectified expressions such as ποίησον ἀγάπην *poíēson agápēn* are intended to re-shape

a message from the speaker considering the addressee's possible reaction to said message, more specifically, to try to make the addressee's response more favourable.

In this section, I have rearranged the polysemy of ποιέω ἀγάπην *poiéō agápēn* according to semantic, syntactic, and pragmatic criteria. Despite the variation that it displays, ποιέω ἀγάπην *poiéō agápēn* has in fact only three basic values: support-verb construction, realisative construction, and pragmatic marker. The biggest polysemy can be found within the group of realisative constructions. However, the meanings of all the constructions discussed can be reduced to "to carry out a [specific kind of] act of love". This could help to simplify and re-organise the lexicographical entries of ποιέω ἀγάπην *poiéō agápēn* across dictionaries.

In the following section, I focus on the evidence for the pragmaticalisation of ποίησον ἀγάπην *poíēson agápēn* in the cases of (vii) and (viii). I address the morphological and syntactic parameters that prove a higher or lower degree of pragmaticalisation and explain the specific pragmatic contexts in which ποίησον ἀγάπην *poíēson agápēn* is used.

4. Syntactic changes

Together with the semantic development of ποιέω ἀγάπην *poiéō agápēn* explored in Section 3, some syntactic changes became apparent that point to its progressive acquisition of pragmatic properties (see Section 2). Below, I explore the degree of ossification of the categories of number, word order, and syntactic integration of ποίησον ἀγάπην *poíēson agápēn* as a pragmatic marker. Lastly, I describe the pragmatic contexts in which it is used.

4.1 Morphological number and number of addressees

The first hypothesis that I tested was whether there were phrases where number was used inconsistently with the context in which they are used, namely texts in which ποίησον ἀγάπην *poíēson agápēn* occurs in the singular with a plural addressee or texts where ποιήσατε ἀγάπην *poiḗsate agápēn* in the plural occurs with a singular addressee. If the phrase ποίησον ἀγάπην *poíēson agápēn* were fully pragmaticalised, it could also be used to address a plural addressee. However, no such incoherences were found, which speaks for a lower degree of pragmaticalisation. When the singular is used, it addresses a single interlocutor, (24), whereas the plural is used when speaking to groups of two or more people, (25).

(24) Dorotheus, *Vita sancti Dosithei*, 6.9

ποίησον ἀγάπην, *κῦρι*, *μάθε* τί ἔχει οὗτος ὁ ἀδελφός, ὅτι κλαίει, καὶ οὐκ οἴδαμεν διὰ τί

poíēson agápēn, *kȳri, mátʰe tí ékʰei ʰoûtos ʰo adelpʰós, ʰóti klaíei, kaì ouk oídamen dià tí*

"**Please**, *lord, find out* what afflicts this brother, for he is crying and we do not know why."

(25) Isaias, *Asceticon (e cod. Hierosol. 109)*, 5.1.8

ἐὰν δὲ εἴπῃ, **ποιήσατε ἀγάπην**, *διδάξατέ* με, ὅτι οὐκ οἶδα…

*eàn dè eípēi, **poiḗsate agápēn**, didáksaté me, ʰóti ouk oîda…*

"If he says: '**Please**, *teach* (pl.) me, for I do not know…'"

In both examples there are explicit indicators of the number of interlocutors. In (24) there is a singular vocative κῦρι *kȳri* "lord" followed by a singular imperative μάθε *mátʰe* "learn". In (25), there is another plural imperative διδάξατε *didáksate* "teach."

However, ποίησον ἀγάπην *poíēson agápēn* in the singular is used a lot more frequently than ποιήσατε ἀγάπην *poiḗsate agápēn* in the plural, with 108 instances *vis à vis* only 14. This partially delimits the context of use of ποίησον ἀγάπην *poíēson agápēn*: it is mostly used in one-to-one conversation, rather than addressing groups of people.

4.2 Syntactic position

Despite being a longer phrase, ποίησον ἀγάπην *poíēson agápēn* behaves similarly to other imperative particles, such as φέρε *pʰére*, ἄγε *áge*, ἴθι *ítʰi* (≈ "come on!"), etc. Zakowski (2018: 64) indicates that these imperatives appear preferentially in the first position of the sentence, which he links to their acquisition of pragmatic meaning. La Roi (2022: 182) counter-argues that this correlation is not to be taken lightly, since another imperative, ἀμέλει *amélei* "don't worry", gradually acquires a higher syntactic mobility as it recruits intersubjective functions. In short, different values of ἀμέλει *amélei* tend to take different syntactic positions: where non-intersubjective uses of ἀμέλει *amélei* mostly take a first position, intersubjective uses show a higher degree of positional mobility.

Another issue that La Roi (2022: 180, n. 31) raises is the difficulty to define first, second, and other syntactic positions in Greek. For instance, there may be several first or second positions, depending on how we define a position. In this chapter, I understand that the first position of a sentence is its first

prosodic item. If ποίησον ἀγάπην *poíēson agápēn* is preceded by a particle such as ἀλλά *allá*, then it is not in the first position, (26). However, if ποίησον ἀγάπην *poíēson agápēn* is the first item of a sentence in direct speech, no matter how it is inserted into the text, it takes the first position, (27). In this chapter, I only look into whether ποίησον ἀγάπην *poíēson agápēn* is used in the first position or any other position (see Table 6.1).

(26) Barsanuphius et Joannes, *Quaestiones et responsiones ad coenobitas*, 130.26

ἀλλὰ **ποίησον ἀγάπην** καὶ λαβὲ μικρὸν ἐξ ἐκείνου τοῦ οὐρανίου πυρὸς οὗ ἦλθε βαλεῖν ἐπὶ τὴν γῆν ὁ τῶν ἁπάντων Δεσπότης, ἵνα ὅταν σπείρωνται τὰ ὑπὸ τοῦ ἐχθροῦ ζιζάνια, φλέξῃ τὸ πῦρ καὶ καταφάγῃ αὐτά

allà ***poíēson agápēn*** *kaì labè mikròn eks ekeínou tou ouraníou pyròs ʰoû êltʰe baleîn epì tèn gên ʰo tôn ʰapántōn Despótēs, ʰína ʰótan speírōntai tà ʰypò tou ekʰtʰrôu zizánia, pʰléksēi tò pŷr kaì katapʰágēi autá*

"*But* **please/do me a favour** and take a little of that heavenly fire which the Lord of all came to cast upon the earth, so that, when the tares of the enemy are scattered, the fire may burn and consume them."

(27) Joannes Moschus, *Pratum spirituale*, 69.2921.12

ὁ αὐτὸς ἀββᾶς Παλλάδιος, ἐρωτηθεὶς ὑφ' ἡμῶν, *ὅτι* **Ποίησον ἀγάπην**, Πάτερ, εἶπον ἡμῖν ποθεν καὶ ἐκ ποίων λογισμῶν ἦλθες εἰς τὸ μοναχικόν

ʰo autòs abbâs Palládios, erōtētʰeìs ʰyp' hēmôn, ʰóti ***Poíēson agápēn,*** *Páter, eîpon ʰēmîn pótʰen kaì ek poíōn logismôn êltʰes eis tò monakʰikón*

"Abbot Palladius himself, when asked by us, *said*: '**Please**, Father, tell us whence and by which arguments you came to monastic life.'"

There is another circumstance that complicates the matter, which is ποίησον ἀγάπην *poíēson agápēn* being constituted by two separate lexical items, the verb and the noun, which are not fully syntactically incorporated. There is a certain degree of ossification (see La Roi 2022), but mobility of its components is evident in some of the examples. Sometimes, the verb and the noun are separated by a particle, (28), or vocative, (29):

(28) *Apophthegmata patrum (collectio systematica)*, 6.8.27

ποίησον *οὖν* **ἀγάπην** καὶ κήρυξον ἐν τῇ ἐνορίᾳ, μή τις ἀπώλεσεν αὐτό

poíēson *oûn* ***agápēn*** *kaì kḗrykson en têi enoríāi, mḗ tis apṓlesen autó*

"**Please/do me a favour** and announce it in the diocese, lest anyone has lost it [a coin]."

(29) *Apophthegmata patrum (collectio systematica)*, 15.129.43[13]

ποίησον, *ἄδελφε*, **ἀγάπην** καὶ μηδενὶ εἴπῃς ὅτι ἀσθενῶ

***poíēson**, ádelp*^{h}*e*, ***agápēn** kaì mēdenì eípēis* ^{h}*óti ast*^{h}*enô*

"**Please/do me a favour**, *brother*, and don't tell anyone that I am sick."

In these cases, I have tagged ποίησον ἀγάπην *poíēson agápēn* as occupying the first position, since it is the first prosodic item of the sentence, despite its syntactic discontinuity. This is how I have obtained the results that ποίησον ἀγάπην *poíēson agápēn* appears 95 times in a first position, and only 27 in a second or other position (see Table 1). However, the fact Wackernagel's law still applies in some cases indicates that the process of *chunking* has not fully been completed (see Bybee 2015: 124). This aligns with La Roi's (2022: 166) statement that

> the fact that the imperative particles do not reveal the effects of all changes associated with grammaticalization (Zakowski 2018: 63-65) is not problematic, since grammaticalizing constructions do not *need* to undergo every change associated with grammaticalization (e.g. phonetic reduction).

The development of this more complex pragmatic imperative supports the notion that not all the changes associated with grammaticalisation need to be completed for a linguistic item to progressively acquire new pragmatic functions. The contexts in which ποίησον ἀγάπην *poíēson agápēn* is coordinated by means of καί *kaí* "and" with another verb phrase, (21), (23), and (26), show a lower degree of grammaticalisation, since the implication is that ποίησον ἀγάπην *poíēson agápēn* still retains some of its properties as an independent imperative support-verb construction, despite its acquisition of certain pragmatic functions (politeness, attenuation of illocutive force).

4.3 Syntactic integration

Since the phrase ποίησον ἀγάπην *poíēson agápēn* is only partially pragmaticalised, it is possible to trace in the texts the process of syntactic detachment that accompanies its acquisition of pragmatic properties. For example, in older texts, it is used as the main verb of the clause, still with the full meaning of "do me a favour" and taking other complements, (30). This represents only 6 examples out of the 122 that have been analysed.

13 By contrast, see the word order ποίησον ἀγάπην ἄδελφε *poíēson agápēn, ádelphe* in (32). In both cases, the phrase ποίησον ἀγάπην *poíēson agápēn* is coordinated with the other imperative, which implies that the phenomenon of chunking is not linked to the degree of syntactic integration of the phrase.

(30) Daniel Scetiota, *De mendico caeco*, 8 (=19)

ποίησον εἰς ἐμὲ **ἀγάπην**, ἀδελφέ, ὅτι οὐκ ἔχω πόθεν ἀγοράσαι μοι βαΐα τοῦ καμεῖν καὶ τραφῆναι

***poíēson** eis emè **agápēn**, adelp^{h}é, ^{h}óti ouk ék^{h}ō pót^{h}en agorásai moi baΐa toû kameîn kaì trap^{h}ênai*

"[A blind man begs to a passer-by:] **Do me a favour**, brother, for I do not have the means to buy a rod to work and support myself."

There is another text that exemplifies the middle ground between the stage in (30) and the detachment of ποίησον ἀγάπην *poíēson agápēn* by means of coordination with καί *kaí* "and", which is the introduction of an infinitive clause that expands on the meaning of ἀγάπη *agápē*, (31):

(31) Ephraem Syrus, *Sermones paraenetici ad monachos Aegypti*, 42.20

καί, ὅτι κέκλικεν ἤδη ἡ ἡμέρα, **ποίησον ἀγάπην** *ὑποδέξαι ἡμᾶς ἔσω τῆς θύρας σου, τὴν νύκτα ταύτην*, μήποτε πλανωμένων ἡμῶν φάγωσιν ἡμᾶς τὰ θηρία

*kaì, ^{h}óti kékliken édē ^{h}ē ^{h}ēméra, **poíēson agápēn** ^{h}ypodéksai ^{h}ēmâs ésō tês t^{h}ýras sou, tèn nýkta taútēn, mépote planōménōn ^{h}ēmôn p^{h}ágōsin ^{h}ēmâs tà t^{h}ēría*

"And, since the day has declined, **do** us **the favour** *of receiving us behind your doors tonight*, lest the beasts eat us while we wander."

This example is already similar to those with ποίησον ἀγάπην *poíēson agápēn* in the first position followed by a verb in the imperative, like (32) and (33):

(32) *Apophthegmata patrum (collectio systematica)*, 20.22.17

ποίησον ἀγάπην, ἄδελφε, *καὶ εἰπέ* μοι τὴν πολιτείαν σου

***poíēson agápēn**, ádelp^{h}e, kaì eipé moi tèn politeían sou*

"**Please/do me a favour**, brother, *and tell* me where you come from."

(33) *Apophthegmata patrum (collectio alphabetica)*, 377

ποίησον ἀγάπην· *εἰπέ* μοι τὸ ῥῆμα τοῦτο

***poíēson agápēn**· eipé moi tò ^{h}rêma toûto*

"**Please**, *tell* me this word."

There is a big difference between (32) and (33). While ποίησον ἀγάπην *poíēson agápēn* in (32) forms a clause of its own which is coordinated with the clause of εἰπέ *eipé* by καί *kaí* "and" (62 examples), ποίησον ἀγάπην *poíēson agápēn* in (33) is fully detached from the main clause and is used as a left-periphery satellite (54 examples). While (33) illustrates the use of ποίησον ἀγάπην *poíēson agápēn* "please" as a politeness marker, in the other examples it seems to be in some middle ground between the support-verb construction and the pragmatic

marker. As it acquires pragmatic functions, the syntactic realisations of ποίησον ἀγάπην *poíēson agápēn* are reduced to three: it is the main verb of the clause (mostly in independent sentences with no other components) (6 examples), it is coordinated with (62 examples), or juxtaposed (54 examples) to another imperative or interrogative clause.

4.4 Pragmatic function: Requests and demands

The core pragmatic function of ποίησον ἀγάπην *poíēson agápēn* is the expression of politeness. Whether it precedes a request, (34), or a question, (35), it is used to downplay the illocutive force of the speaker's statement.

(34) Joannes Moschus, *Pratum spirituale*, 79.2929

τί κελεύεις, κύριε; Τότε λέγω αὐτῇ· **Ποίησον ἀγάπην**, ἐλθὲ ἕως ὧδε
tí keleúeis, kýrie? tóte légō autêi· ***poíēson agápēn****, elt*^{h}*è* ^{h}*éōs* ^{h}*ôde*
"'What do you command, lord?' Then I tell her: '**Please**, come here.'"

(35) Joannes Moschus, *Pratum spirituale*, 123.2985 (= 22)

ποίησον ἀγάπην, καλόγηρε, πόθεν με οἶδας;
poíēson agápēn*, kalógēre, pót*^{h}*en me oîdas;*
"**Please**, Reverend, whence do you know me?"

However, examples such as (35) in which ποίησον ἀγάπην *poíēson agápēn* introduces a direct question are more obviously pragmaticalised than examples such as (34), in which the distinction between the support-verb construction and the pragmatic marker is blurrier. In a way, (34) is rather similar to those in which ποίησον ἀγάπην *poíēson agápēn* retains its full meaning "to do a favour" and introduces a description of the favour with an infinitive clause, (31), or a coordinated καί *kaí* "and"-clause, (32). It might even have been possible to paraphrase (34) as ποίησον ἀγάπην *ἐλθεῖν poíēson agápēn elt*^{h}*eîn*. But this is not the case with (35), where ποίησον ἀγάπην *poíēson agápēn* is neither syntactically nor semantically linked to the interrogative clause. It seems clear that ποίησον ἀγάπην *poíēson agápēn* in (35) does not convey any meaning related to a favour, but that it is used to politely catch the attention of the addressee and prepare them for a positive outcome of the interaction, whose purpose is to obtain a specific piece of information. Out of the 20 examples of ποίησον ἀγάπην *poíēson agápēn* introducing a question, thirteen are juxtaposed, whereas only six are coordinated by καί *kaí* "and". The examples that show juxtaposition introduce direct questions, (35), whereas the examples that show coordination introduce indirect questions, (23). Furthermore, the use of ποίησον ἀγάπην *poíēson agápēn* introducing questions is chronologically restricted, since most examples appear

in the sixth century CE (see Table 6.1). From the chronological distribution of the examples, it can be stated that there is not a chronological directionality from ποίησον ἀγάπην *poíēson agápēn* + request to ποίησον ἀγάπην *poíēson agápēn* + question, but rather a functional specialisation of the phrase. However, from a sociocultural perspective, pragmaticalised ποίησον ἀγάπην *poíēson agápēn* is restricted to Christian texts, both in content and origin, and is particularly common in hagiographical texts.

5. Conclusions

In this chapter, I have discussed the pragmaticalisation of the phrase ποίησον ἀγάπην *poíēson agápēn* according to three research questions: (i) What is the meaning of ποιέω ἀγάπην *poiéō agápēn*, (ii) what are the indicators of the pragmaticalisation of ποίησον ἀγάπην *poíēson agápēn*, and (iii) to what extent is ποίησον ἀγάπην *poíēson agápēn* a pragmaticalised phrase? First, I have presented an overview of the 122 examples of semi-pragmaticalised ποίησον / ποιήσετε ἀγάπην *poíēson*/*poiḗsete agápēn* (see Table 6.1), which show that the expression peaked in use by the fourth to sixth centuries CE, in Christian and more specifically in hagiographical texts, and declined afterwards, even though it did not completely disappear. Secondly, I have explored the different meanings with which ποιέω ἀγάπην *poiéō agápēn* is used in the *Thesaurus Linguae Graecae*, where I have identified six unpragmaticalised meanings, the first one of which is a support-verb construction, while the rest are realisative constructions (i. "to love", ii. "to perform an act of love", iii. "to make love", iv. "to make peace", v. "to celebrate a banquet", vi. "to do a favour") and two pragmatic uses (vii. "please" followed by a request, viii. "please" followed by a question).

Next, I have studied the pragmatic uses (vii and viii). I have found that morphological number is coherent with the number of addressees of the request or question, which is an indication that the pragmaticalisation of ποίησον ἀγάπην *poíēson agápēn* is only partial. Then, I have discussed the position of ποίησον ἀγάπην *poíēson agápēn* in the left periphery of the clause, which is the typical position for the codification of intersubjective information. Regarding the syntactic integration of ποίησον ἀγάπην *poíēson agápēn* in the clause, coordination with καί *kaí* "and" is the most common form of linking ποίησον ἀγάπην *poíēson agápēn* to the main clause. However, καί *kaí* "and" is used to link another imperative, which means that direct questions may only be introduced by ποίησον ἀγάπην *poíēson agápēn* without καί *kaí* "and". The last conclusion permeates all the points of the discussion: ποίησον ἀγάπην *poíēson agápēn* as a pragmatic function is only used in two contexts within Christian literature, one of which is the introduction of requests and the other is the introduction of questions. In both cases, its function is that of catching the addressee's attention,

softening the illocutive force of the statement, and predisposing the addressee to a positive outcome of the interaction.

Acknowledgements

This chapter has been made possible thanks to the funding of Interacción del léxico y la sintaxis en griego antiguo y latín 2: Diccionario de Colocaciones Latinas (DiCoLat) y Diccionario de Colocaciones del Griego Antiguo (DiCoGra) (PID2021-125076NB-C42) and the Spanish Ministry of Science, Innovation and Universities (PRE2022-102525). I thank M.D. Jiménez López, J.M. Baños, the anonymous reviewers and the editor of this volume for their valuable insights on a previous version of this chapter.

Bibliography

Alonso Ramos, Margarita. 2004. *Las construcciones con verbo de apoyo*. Madrid: Visor Libros.

Baños, José Miguel. 2014. Construcciones con verbo soporte, extensiones y elecciones preferentes: *bellum parare*, *suscipere*, *trahere* y *conficere. BOSELAT* 1. 5–11.

Baños, José Miguel, Cristina Tur, María Dolores Jiménez López, and María Isabel Jiménez Martínez (eds.). 2022. *Collocations in Theoretical and Applied Linguistics: From Classical Languages to Romance Languages*. Madrid: Guillermo Escolar/SEEC.

Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness. Some Universals in Language Use*. Cambridge: Cambridge University Press.

Bybee, Joan. 2015. *Language Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9781139096768

Caffi, Claudia. 2005. *Mitigation*. Leiden: Brill. https://doi.org/10.1163/9780080466224

Cock, Alwies J. C. M. 1981. ΠΟΙΕΙΣΘΑΙ: ΠΟΙΕΙΝ. Sur les critères déterminant le choix entre l'actif ΠΟΙΕΙΝ et le moyen ΠΟΙΕΙΣΘΑΙ. *Mnemosyne* 34 (1). 1–62.

Degand, Liesbeth, and Jacqueline Evers-Vermeul. 2015. Grammaticalization or Pragmaticalization of Discourse Markers? More than a Terminological Issue. *Journal of Historical Pragmatics* 1(16). 59–85. https://doi.org/10.1075/jhp.16.1.03deg

Denizot, Camille. 2022. How to Be Polite without Saying "Please" in Classical Greek?: The Role of δή in Polite Requests. In Luis Unceta Gómez and Łukasz Berger (eds.), *Politeness in Ancient Greek and Latin*, 77–102. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009127271.004

Fendel, Victoria Beatrix. 2022. *Coptic Interference in the Syntax of Greek Letters from Egypt*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780192869173.001.0001

— 2023. "I Haven't Got a Clue!": Assessing Negation in Classical Greek Support-Verb Constructions. *Journal of Greek Linguistics*. 23(2). 139–63. https://doi.org/10.1163/15699846-02302004

— 2024. Celebrating Diversity: The Origins and Pathways of Three Support-verb

Constructions. *Lexis* 24. https://doi.org/10.4000/12cvp

— 2025. *Giving Gifts and Doing Favours: Literary Classical Attic Greek Support-Verb Constructions*. Leiden: Brill. https://doi.org/10.1163/9789004730526

Gross, Gaston. 1999. Verbes supports et conjugaison nominale. *Revue d'études Francophones* 9. 70–92.

— 2004. Pour un bescherelle des prédicats nominaux. *Lingvisticae Investigationes* 27(2). 343–358. https://doi.org/10.1075/li.27.2.14gro

Goutsos, Dionysis. 2017. A Corpus-based Approach to Functional Markers in Greek: Exploring the Role of Position. In Chiara Fedriani and Andrea Sansò (eds.), *Pragmatic Markers, Discourse Markers and Modal Particles: New Perspectives*. 125–150. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.186.05gou

Holton, David, Geoffrey Horrocks, Marjolijne Janssen, Tina Lendari, Io Manolessou, and Notis Toufexis. 2019. *The Cambridge Grammar of Medieval and Early Modern Greek*. 4 vols. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316632840

Horrocks, Geoffrey. 2014. High Register Byzantine Greek: Diglossia and What Lay Behind It. In Caterina Carpinato and Olga Tribulato (eds.), *Storia e Storie Della Lingua Greca*, 49–72. Venice: Edizioni Ca' Foscari.

Jiménez López, María Dolores. 2016. On Support Verb Constructions in Ancient Greek. *Archivio Glottologico Italiano* 101(2). 180–204.

— 2021. Γίγνομαι as the lexical passive of the support verb ποιέω in Ancient Greek. In Georgios K. Giannakis, Luz Conti, Jesús de la Villa, and Raquel Fornieles (eds.), *Synchrony and Diachrony of Ancient Greek. Language, Linguistics and Philology*, 227–240. Berlin: De Gruyter. https://doi.org/10.1515/9783110719192-018

La Roi, Ezra. 2022. Weaving Together the Diverse Threads of Category Change: Intersubjective ἀμέλει "of course!" and Imperative Particles in Ancient Greek. *Diachronica* 39(2). 151–192. https://doi.org/10.1075/dia.20031.lar

— 2025. Van smeken naar "alsjeblieft" in het Oudgrieks: De historische pragmatiek van performatieve taalhandelingen. *Lampas* 58(4). 359–381. https://doi.org/10.5117/lam2025.4.005.laro

Mel'čuk, Igor. 2023. *General Phraseology: Theory and Practice*. Amsterdam: John Benjamins. https://doi.org/10.1075/lis.36

Thayer, Joseph Henry. 1887. *Greek-English Lexicon of the New Testament, Being Grimm's Wilke's Clavis Novi Testamenti*. New York: Harper & Brothers.

Toufexis, Notis. 2008. Diglossia and Register Variation in Medieval Greek. *Byzantine and Modern Greek Studies* 32. 203–17. https://doi.org/10.1179/174962508x322687

Vivès, Robert. 1984. L'aspect dans les constructions nominales prédicatives: "avoir", "prendre", verbe support et extension aspectuelle. *Lingvisticae Investigationes* 8(1). 161–185.

Zakowski, Samuel. 2018. The Evolution of the Ancient Greek Deverbal Pragmatic Markers áge, íthi and phére. *Journal of Historical Pragmatics* 19(1). 55–91. https://doi.org/10.1075/jhp.16009.zak

General abbreviations

Bailly = Bailly, Anatole. 2020. *Bailly 2020 Hugo Chávez. Dictionnaire grec-français.* With Antoine Boquet, Benjamin Georges, Gréco Gérard, André Charbonnet, Mark De Wilde, and Bernard Maréchal. https://bailly.app/

DGE = Adrados, Francisco Rodríguez, Elvira Gangutia Elícegui, Dolores Lara Nava, et al. 1980. *Diccionario Griego-Español En Línea*. CSIC. http://dge.cchs.csic.es/xdge/

DiCoGra = Jiménez López, María Dolores, and José Miguel Baños. n.d. *DiCoGra: Diccionario de Colocaciones Del Griego Antiguo.* https://dicogra.iatext.ulpgc.es/

LBG = Trapp, Erich. 2020. *Lexikon zur Byzantinischen Gräzität*. Vienna: Austrian Academy of Sciences. https://stephanus.tlg.uci.edu/lbg/

LSJ = Liddell, Henry George, Robert Scott, and Henry Stuart Jones. 1996. *A Greek-English Lexicon.* 9th edn. with revised supplement. Oxford: Clarendon. https://stephanus.tlg.uci.edu/lsj

7. Between word and phrase: The puzzle of Coptic light-verb constructions (2nd to 5th c. CE)

Nina Speransky

1. Introduction[1]

Coptic represents the final documented stage of the Egyptian language, an independent branch of the Afroasiatic family that was spoken in the northern Nile Valley. It is attested from around 300 CE (Richter 2023),[2] and remained in use as a living language of written communication until the eleventh century CE, when it was gradually replaced by Arabic following the Arab conquest of Egypt in the seventh century CE. Today, it survives as a liturgical language of the Coptic Orthodox Church.

Early Coptic already exhibits substantial dialectal variation (Richter 2023). Of these dialects, Sahidic (southern) and Bohairic (northern) are by far the best attested. Sahidic alone accounts for ca. 90% of the Coptic textual corpus. For this reason, the present study primarily draws on Sahidic data, occasionally comparing Sahidic and Bohairic light-verb constructions (LVCs henceforth). Examples are marked with the conventional sigla, (*S*) for Sahidic and (*B*) for Bohairic.

The issue of LVCs is briefly treated in two recent descriptive grammars of Sahidic (Reintges 2018; Layton 2011).[3] The present chapter builds on their data, but attempts to situate Coptic LVCs within general typology. Addressing the

1 Dataset: https://doi.org/10.5281/zenodo.20330827. The data was collected manually by inspecting the text. The glossing conventions here are in accordance with the Leipzig Glossing Rules, and the transliteration follows the Leipzig-Jerusalem transliteration system (Grossman and Haspelmath 2014).

2 The Greek-based script Coptic—supplemented by several Demotic signs—likely developed earlier, around 100 CE (Fendel 2022).

3 In Layton (2011), LVCs are treated as “compound verbs”.

©2026 Nina Speransky, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.07

inventory of LVCs in both dialects and investigating their argument licensing strategies, the chapter shows that they function primarily as a derivational strategy while remaining accessible to syntactic operations such as agreement and indexing.

Unless otherwise specified, the data for the study are drawn from the Coptic Scriptorium corpus (https://data.copticscriptorium.org), a digitally annotated collection of Coptic texts. The translations are mine. Given the canonical nature of the biblical material that resists any changes and the uncertainties surrounding the dating of individual manuscripts, the corpus does not allow for a detailed diachronic investigation. The present study therefore adopts a primarily synchronic perspective.

The chapter is organised as follows. Section 1 argues that not all combinations of bound verb forms with nominal objects can be analysed as noun incorporation or lexical affixation and proposes a set of criteria correlating with LVCs. Section 2 presents a survey of the inventory of LVCs in Sahidic and Bohairic. Section 3 examines argument assignment strategies associated with LVCs. Section 4 addresses the relationship between LVCs and their synthetic counterparts arguing that attributing a uniform grammatical function (such as telicity marking) to Coptic LVCs is implausible. Finally, Section 5 considers the implications of the Coptic data for the Lexical Integrity Principle *pace* Booij (2009).

2. Separating light-verb constructions from morphologically bound complementation

The elementary units of Coptic are not words—the notion of "word" can hardly be meaningfully defined for Coptic (Haspelmath 2015)—but rather stress groups, also known as "bound groups" (Layton 2011). These groups mostly consist of a nominal or verbal root preceded by a series of phonologically dependent grammatical morphemes. Synthetic constructions with postposed person and tense-aspect markers (TAM henceforth), characteristic of earlier stages of Egyptian, underwent periphrasis in a pre-Coptic stage, followed by secondary synthesis in Coptic. This development resulted in structures with preposed person and TAMs (Haspelmath 2015).

Coptic employs two verbal clause patterns: the TAM–subject–verb eventive pattern, and the subject–predicate durative pattern. In the eventive pattern, the verb may appear either in a free form (the so-called *status absolutus*) or in a bound form. Transitive[4] verbs typically exhibit two bound forms: one used

4 As opposed to the definition in Layton (2011), the term here does not refer to any bivalent verb, but rather the valency class corresponding to the "kill" or "hit" valency pattern. In Coptic, this

with pronominal objects (*status pronominalis*), and another used with nominal objects (*status constructus*). For instance, ⲉⲓⲣⲉ *eire* "do" has a pre-pronominal form ⲁⲁ *aa*= and a pre-nominal form ⲣ *r*- (1).

(1a) Genesis 30:30

ⲉⲕⲱⲁⲛⲉⲓⲣⲉ	ⲙⲡⲉⲓⲱⲁϫⲉ
e=k-šan-eire	*m-pei-šače*
if.2SGM-do	ACC-DEM.SGM-word

"If you do this"

(1b) Numbers 22:28

ⲟⲩ	ⲡⲉⲛⲧⲁⲓⲁⲁϥ	ⲛⲁⲕ
ou	*p-ent-a=i-aa=f*	*na=k*
what	DEF.SGM-PST.REL-PST-1SG-do-3SGM	DAT-2SGM

"What have I done to you?"

(1c) Exodus 1:20

ⲡⲛⲟⲩⲧⲉ	ⲇⲉ	ⲁϥⲣⲡⲡⲉⲧⲛⲁⲛⲟⲩϥ	ⲛⲙⲙⲉⲥⲓⲟ
p-noute	*de*	*a=f=r-p-petnanouf*	*n-m-mesio*
DEF.SGM-god	PTCL	PST-3SGM-do-DEF.SGM-good	DAT-PL-midwife

"And God did well to the midwives."

The combination of *status constructus* forms with nominal objects in Coptic remains in need of a morphosyntactic interpretation that is meaningful from a typological perspective. A morphological merger of the bound verbal form with an internal object noun may resemble two distinct phenomena: noun incorporation understood as a verbal compound of a nominal and a verbal stem (Mithun 1984) and lexical affixation, whereby bound morphemes with root-like meaning attach to host (Mithun 1997). Since lexical affixes are regarded as a step in the evolution of noun incorporation (Mithun 1997: 366), the difference between the two constructions amounts to their semantics. In incorporation, the verb is semantically specified by the incorporated noun, while in lexical affixation, the affix typically contributes a highly generalised meaning.

One could argue that Coptic exhibits verb–noun constructions of both types. The combinations ⲥⲕ-ⲙⲟⲟⲩ *sk-moou* "to draw water" and ⲣⲉⲧ-ⲧⲛϩ *ret-tnh* "to grow wings" in (2) and (3) resemble Mithun's type I of noun incorporation, that

class reveals additional symptoms of transitivity, such as sensitivity to the predicate aspect or the referentiality status of the object (see Speransky 2022).

is lexical compounding, where a verb stem and a juxtaposed unmodified non-referential noun together denote one activity. This interpretation is supported by the fact that each combination corresponds to a single verbal lexeme in the Greek source texts.

(2) Sahidic, Exodus 2:17

ⲁϥⲥⲕⲙⲟⲟⲩ	ⲛⲁⲩ	ⲁϥⲧⲥⲟ
a-f-sk-moou	*na=u*	*a-f-tso*
PST-3SGM-draw-water	DAT-3PL	PST-3SGM-give_drink

ⲛⲛⲉⲩⲉⲥⲟⲟⲩ
n-neu-esoou
ACC-POSS.PL:3PL-sheep

“He drew water (ἀντλέω) for them, and gave their sheep to drink.”

(3) Sahidic, Isaiah 40:31

ⲥⲉⲛⲁⲣⲉⲧⲧⲛϩ	ⲛⲑⲉ	ⲛϩⲉⲛⲁϩⲱⲙⲉ
se-na-ret-tnh	*n-t-he*	*n-hen-ahôme*
3PL-FUT-grow-wings	in-DEF.SGF-way	GEN-PL.INDEF-eagle

“They will grow wings (πτεροφυέω) like eagles.”

Conversely, some accounts (for example, Grossman 2023) propose interpreting bound verbs in bound verb-noun combinations as lexical affixes, based on their strict adjacency to the nominal object, the syntactic contiguity of the combination (notably, its resistance to the insertion of Wackernagel clitics), and the alleged association of the object noun with the lower end of the referentiality scale.

However, neither noun incorporation nor lexical affixation provides a fully adequate model for the Coptic data. Each analysis captures certain properties of Coptic verb–noun combinations while failing to account for others. First, in these combinations the noun occupies the usual position of a nominal argument; this does not comply with the strict definition of noun incorporation proposed by Haspelmath (2025).[5] Further, in canonical incorporation constructions, nouns are typically non-referential or unmarked for specificity and definiteness (Mithun 1984: 849ff.). Similar restrictions apply to lexical affixes, with additional diagnostics: lexical affixes form a closed class (Mithun 1997: 369), and verbal prefixes undergoing lexical affixation are subject to significant semantic bleaching, combining with another root to form a word expressing a single

5 According to this definition, noun incorporation is a compound in which “the noun […] occurs in a position where nominal patient arguments cannot occur” (Haspelmath 2025: 2).

primary concept. Coptic construct verb phrases do not consistently conform to these diagnostic criteria.

a) Verbs exhibiting construct forms do not constitute a closed lexical class. On the contrary, nearly all transitive verbs in the native verbal inventory possess a construct form that cliticises to nominal objects.

b) The construction is not restricted to low-referentiality nouns. Examples (4) to (7), which are by no means marginal, demonstrate that the object position may be occupied by definite nouns with possessive modifiers, demonstrative pronouns, or even proper names. Rather than correlating with low referentiality, the construction is associated with low informational status (Speransky 2022),[6] a property compatible with the third type of noun incorporation identified by Mithun (1984).

c) With referential nouns, construct verb phrases retain compositional semantics and do not merge into a single lexical concept, as illustrated in (4) and (5):

(4) Sahidic, Judges 3:20

ⲁϥⲥⲟⲩⲧⲛⲧⲉϥϭⲓⲍ					ⲛϩⲃⲟⲩⲣ	ⲉⲃⲟⲗ
a-	*f-*	*soutn-*	*tef-*	*cič*	*n-hbour*	*ebol*
PST-	3SGM-	straighten-	POSS.SG	F:3SGM-hand	ATTR-left	out

“He prostrated his left hand”

(5) Sahidic, Pistis Sophia AQ8

ⲁⲡⲥⲱⲧⲏⲣ	ϭⲗⲡ	ⲛⲁⲓ	ⲛⲁⲛ
a-p-sôtêr	*clp*	*nai*	*na=n*
PST-DEF.SGM saviour	reveal	this.PL	DAT-1PL

“The Saviour revealed these (things) to us”

d) Referential nouns preserve their syntactic independence: they may be modified by possessive prefixes, as in (4), or by relative clauses.

e) When the object is highly referential (for example, a demonstrative or proper name), the two elements of the combination clearly do not form a lexical unit at the formal level. Otherwise, one would be forced to posit lexemes such as *straighten-his-hand*, *reveal-these*, or even *confuse-Lot* and *send-Moses*, as in (6) and (7), or even assume that an entire clause may host a lexical affix, as in (8) where the construct form *cn* of the verb *cine* “to find” precedes a clausal complement.

6 But cf. Engsheden (2008: 341), who argues the opposite.

(6) Sahidic, Genesis 19:15

ⲁⲛⲁⲅⲅⲉⲗⲟⲥ			ⲯⲧⲣⲧⲣⲗⲱⲧ	
a-	*n-*	*aggelos*	*štrtr-*	*lôt*
PST-	DEF.PL	angel	confuse	Lot

“The angels confused Lot.”

(7) Sahidic, Psalms 104:26

ⲁⲫⲭⲉⲩⲙⲱⲩⲥⲏⲥ				ⲡⲉⲫⲍⲙⲍⲁⲗ	
a-	*f-*	*čeu*	*môusês*	*pef-*	*hmhal*
PST	3SGM	send	Moses	POSS.SGM:3SGM	servant

“He sent Moses his servant.”

(8) Sahidic, Luke 19:48

ⲁⲩⲱ	ⲙⲡⲟⲩⲟⲛ			ⲛⲉⲩⲛⲁⲣ			ⲟⲩ	ⲛⲁⲫ
auô	*mp-*	*ou-*	*cn*	*ne-*	*u-*	*na-r*	*ou*	*na=f*
And	NEG.PST-	3PL-	find	PRET-3P	L-	FUT-do	what	DAT-3SGM

“And they could not find what to do to him.”

To sum up, none of the usual diagnostics for LVCs, such as membership of the verb in a closed class, its semantic bleaching, or non-referentiality of the complement noun, consistently applies to the whole class of Coptic construct predicates. Hence, treating all the instantiations of bound verb forms under the umbrella terms “noun incorporation” or “lexical affixes” seems to misrepresent the situation in Coptic.

However, there exists in Coptic a class of construct verb phrases easily recognizable by the sheer frequency of specific verb-noun collocations, where the verbs are best analysed as light verbs or even as affixes. However, the concept of a light verb has to be modified for Coptic. The usual, more or less intuitive idea of a light verb is based, in most general terms, on the division of labour between the verb and the complement noun: the verb carries the grammar of the combination (tense and agreement in Jespersen (1954), aspect in Sobolak (2023), etc.), whereas the noun defines the semantics of the combination. In Coptic, this concept only applies at the semantic level, since the tense-aspect-mood and person inflection is the property of a clause pattern as a whole (see above) and is not marked on the verb, anyway. Therefore, the essential criterion identifying this class appears to be a semantically light verb combined with an inherently (as with *nomina actionis*) or contextually non-referential object noun. The criterion of verb

semantic bleaching permits us to distinguish between light-verb phrases and noun incorporations where a non-referential noun follows a full lexical verb in the construct form, such as ⲥⲕ-ⲙⲟⲟⲩ *sk-moou* “to draw water” (ⲥⲱⲕ *sôk* “to draw”, ⲙⲟⲟⲩ *moou* “water”), ⲥⲉⲡⲧⲃⲧ *ceptbt* “to fish” (ⲥⲱⲡ *côp* “to catch”, ⲧⲃⲧ *tbt* “fish”), ⲙⲉϣⲧⲡⲕⲁϩ *meštpkah* “to search the land, to spy” (ⲙⲟⲩϣⲧ *moušt* “to watch”, ⲕⲁϩ *kah* “land”), ⲟⲩⲉϩⲥⲁϩⲛⲉ *ouehsahne* “to order” (ⲟⲩⲱϩ *ouôh* “to set”, ⲥⲁϩⲛⲉ *sahne* “to order”). However, this criterion is imprecise as the desemantisation of the verb is gradual, and phrases with non-referential nouns form a continuum in this respect. For instance, the combinations of ⲣ- *r-* “to do” with a noun, such as ⲣ-ϩⲱⲃ *r-hôb* “to work”, ⲣ-ϣⲡⲏⲣⲉ *r-spêre* “to wonder” should probably be considered LVCs.

The tight link between the verb and the noun can also be manifested in the following:

1. Non-compositionality/idiomaticity of the combination;
2. In cases where the verb-noun phrase takes arguments, the verb-noun phrase can be further expanded: argument licensing by the verb-noun combination as a whole, rather than by the verb or the noun alone; in the extreme case, the ability to take an argument identical to the object noun;
3. Non-modifiable object noun.

Non-modifiability of the object noun is not a decisive criterion for Coptic: several LVCs index their A[gent]- or P[atient]-argument as a possessor on the object noun (for example, ⲡⲉϥ-ⲟⲩⲟⲓ *pef-ouoi* “his course” in (9)). Importantly, such possessive indexing does not render the noun referential: as in English *take one’s leave*, the nominal component does not denote a discourse entity, and the construction still qualifies as an LVC.

(9) Sahidic, Genesis 29:10

ⲁⲓⲁⲕⲱⲃ	ϯⲡⲉϥⲟⲩⲟⲓ	ⲉϩⲟⲩⲛ
a-iakob	*ti-pef=ouoi*	*ehoun*
PST-Jacob-	give-POSS.SGM=3SGM-course	inside

“Jacob approached.”

The constructions examined here only partially align with the cross-linguistic profile of light-verb constructions as defined in annotation frameworks such as those developed within the PARSEME initiative. In particular, the nominal element in Coptic does not denote an action or event and, in most cases, does not license syntactic arguments (see Section 4). Furthermore, Coptic light-verb constructions seem to appear as one (phonological word), also in many

orthographic traditions. Some of the constructions display all the properties typically associated with verbal idioms, including morphological, syntactic, and lexical inflexibility, semantic non-compositionality, and tolerance of so-called cranberry items, that is, combinations with semantically opaque and lexically frozen parts.

An idiomatic construction whose arguments are licensed not by any of its members, but by the construction as a whole is effectively a univerbation.[7] A tell-tale sign of univerbation is the use of identical nouns for both the immediate direct and the oblique objects of the phrase, as *ti-htê=f hm pef-hêt* "to heed in his heart" (literally "to give heart in his heart") in (10) or ⲧ-ⲣⲉⲛ ⲡⲉϥ-ⲣⲁⲛ *ti-ren pef-ran* " to name him by the name" (literally "to give name his name") in (11).

(10) Sahidic, Exodus 9:21

ⲡⲉⲧⲉⲙⲡϥⲧϩⲧϥ			ⲁⲉ	ϩⲙ
p-	*ete-*	*mp-f-ti-htê=f*	*de*	*hm*
ART-	REL-	NEG.PST-3SGM-give-**heart**-3SGM	PTCL	in

ⲡⲉϥϩⲏⲧ	ⲉⲡϣⲁϫⲉ	ⲙⲡϫⲟⲉⲓⲥ
pef-hêt	*e-p-šače*	*m-p-čoeis*
POSS.SGM:3SGM-**heart**	ALL-DEF.SGM- word	GEN-DEF.SGM-lord

"He who did not pay attention in his heart to the word of the Lord…"

(11) Bohairic, Genesis 5:2

ⲁϥⲧⲣⲉⲛ		ⲡⲉϥⲣⲁⲛ	ϫⲉ	ⲁⲁⲁⲙ
a-	*f-ti-ren*	*pef-ran*	*če*	*adam*
PST-	3SGM-give-**name**	POSS.SGM:3SGM-**name**	QUOT	Adam

"He named him Adam."

Given the limited applicability of the notion of word in Coptic grammar, one cannot establish any meaningful distinction between a light verb and a verbalising affix, except in clear cases of univerbation, such as the morphological alternation of a combination as a whole (such cases are discussed in Section 3). Since such morphological alternations constitute an

7 A similar point is made, e.g., in Khachaturyan (2017: 15) for Mano (Mande family), a language genealogically distant from Coptic and belonging to a different macro-areal zone: "…in Mano, just like in Kurdish, the argument sharing approach fails in certain N[oun]+V[erb] combinations …in these cases, the external argument must be assigned only by the combination itself. These combinations should be considered more syntactically and semantically bound than those combinations where the external argument is assigned by the combination's nominal component."

exception, they probably cannot define the entire class. Therefore, though there is no compelling reason to abandon the term light verbs that is used for such constructions in Coptic linguistics (Reintges 2018; Zakrzewska 2017), we should keep in mind that Coptic LVCs clearly represent a transitional stage between a MWE and a compound.

3. The inventory of light-verb constructions and their components

The overall number of the light-verb phrases corresponding to the above criteria in Coptic is difficult to estimate, because various dialects and even various corpora within the same dialect differ considerably in this respect. Thus far, no cross-dialectal description of Coptic light verbs has been undertaken. Even the most authoritative reference work, Crum's dictionary, does not contain all such compounds. It is clear, however, that there are several dozen such forms in the Sahidic and Bohairic biblical corpora, which means that these constructions constitute a notable part of the verbal lexicon. Some of them are common to both dialects: they consist of the same lexical elements, and the contexts of Sahidic constructions consistently match those of Bohairic ones. Examples (Sahidic/Bohairic) include: ⲣ-ⲛⲟⲃⲉ / ⲉⲣ-ⲛⲟⲃⲓ *r-nobe / er-nobi* "to sin", ⲣ-ⲡ-ⲱⲃϣ *r-p-ôbš* "to forget", ⲣ-ϩⲱⲃ *r-hôb* "to work", ⲣ-ϣⲡⲏⲣⲉ / ⲉⲣ-ϣⲫⲏⲣⲓ *r-špêre / er-špʰêri*, "to feel amazement", ⲣ-ⲡ-ⲙⲉⲉⲩⲉ / ⲣ-ⲫ-ⲙⲉⲩⲓ *r-p-meeue / er-pʰ-meui* "to remember", ⲣ-ⲡⲁϩⲣⲉ / ⲉⲣ-ⲫⲁϩⲣⲓ *r-pahre / er-pʰahri* "to cure", ⲣ-ϩⲟⲧⲉ / ⲉⲣ-ϩⲟⲧ *r-hote / er-hotⁱ* "to fear", ϫⲓ-ⲙⲟⲉⲓⲧ / ⲥⲓ-ⲙⲱⲓⲧ *či-moeit / ci-môit* "to lead (take way)", ⲕⲁ-ⲣⲱ / ϧⲁ-ⲣⲱ *ka-rô= / kʰa-rô=* "to keep quiet (hold mouth)", ⲧ-ⲛⲁⲁⲕⲉ / ⲧ-ⲛⲁⲕϩⲓ *tⁱ-naake / tⁱ-nakhi* "to feel birth pains, to give birth", ϫⲓ-ϣⲓⲡⲉ / ⲥⲓ-ϣⲓⲡⲓ *ci-šipe / ci-šipi* "to be ashamed", ⲧ-ⲥⲃⲱ *tⁱ-sbô* "to teach", ⲥⲙ-ⲡ-ϣⲓⲛⲉ / ϫⲉⲙ-ⲡ-ϣⲓⲛⲓ *cm-p-šine / čem-p-šini* "to visit", ⲫⲓ-ⲣⲟⲟⲩϣ / ⲫⲓ-ⲣⲱⲟⲩϣ *fi-roouš / fi-rôouš* "to care". Light-verb derivation can also target Greek-origin nouns, for example, ⲉⲣ-ϩⲉⲗⲡⲓⲥ (B) *er-helpis* "to hope", ⲣ-ϧⲣⲉⲓⲁ / ⲉⲣ-ϧⲣⲓⲁ *r-kʰreia / er-kʰria* "to need".[8] Yet, with loanwords the pattern is evidently less productive.

Other constructions are dialect-specific, their semantic equivalents or translational counterparts in the other dialect being synthetic verbs. Some examples are given in Table 7.1 (Sahidic LVCs) and Table 7.2 (Bohairic LVCs).

8 Interestingly, the cognate Greek verb χράομαι *khraomai* "to yearn after, to need" was borrowed into Coptic as ϧⲣⲱ / ϧⲣⲁⲥⲑⲁⲓ *khrô* / *khrasthai*, but with a semantic shift towards the meaning "to use, to apply": Coptic Dictionary Online, ed. by the Koptische/Coptic Electronic Language and Literature International Alliance (KELLIA), https://coptic-dictionary.org/entry.py?tla=C11004

Sahidic	Bohairic
ϫι-ⲧⲁⲉιο *či-taeio* “to glorify, to give honours”	ϣογϣογ *šoušou* “to boast”
ϫι-ϣⲕⲁⲕ *či-škak* “to cry out”	ωϣ ⲉⲃⲟⲗ *ôš ebol* “to cry out”
ϯ-ογοι *ti-ouoi* “to approach”	ϩωⲛⲧ ⲉϩⲟⲩⲛ *hônt ehoun* “to approach” ι ⲉ- / ϩⲁ-/ ϣⲁ-, *i e-/ha-/ša-* “to go” ϣⲉ ϣⲁ- *še ša-* “to go” ⲙⲱϣι *môši* “to go” ϭⲟϫι *coči* “to run”
ϣⲡ-ⲧⲟⲟⲧ= *šp-toot=* “to betroth”	ωⲡ ⲛⲥⲱ= ⲛ- *ôp nsô= n-* “to count”
ϫι-ϩⲣⲁ *či-hra=* “to contemplate, to think about”	ⲁⲣⲉϩ *areh* “to guard” ⲥⲁϫι *sači* “to speak” ⲉⲣ-ⲙⲉⲗⲉⲧⲁⲛ *er-meletan* “to care”

Table 7.1 Sahidic LVCs vs. Bohairic synthetic counterparts in the Biblical corpus.[9]

Bohairic	Sahidic
ϭι-ϕει *ci-fei* “to jump”	ϥωϭⲉ *fôce* “to jump”
ϯ-ⲣⲉⲛ *ti-ren* “to call (by name)”	ⲙⲟⲩⲧⲉ *moute* “to call (by name)”
ϭι-ϣϣωογ *ci-ššôou* “to desire”	ⲉⲡⲓⲑⲩⲙⲉⲓ *epithymei* “to desire” ⲙⲉ *me* “to love” ⲟⲩⲱϣ *ouôš* “to want”
ⲉⲣ-ϭι-ⲥⲁⲛⲓⲥ *er-ci-sanis* “to doubt”	ⲙⲉⲉⲩⲉ *meeue*[10] “to think”
ⲉⲣ-ⲟⲩⲱ *er-ouô* “to answer”	ϣⲁϫⲉ *šače* “to speak” ⲟⲩⲱϣⲃ *ouôšb* “to answer” ⲁⲡⲟⲗⲟⲅⲓⲍⲉ *apologize* “to answer”
ϯ-ⲙⲁϯ *ti-mati* “to fit, to agree, to suit”	ⲫⲓ ⲙⲛ ⲛⲉ=ⲉⲣⲏⲩ *fi mn ne=erêu* “to fit, to agree, to suit” ⲟⲩⲱϣ *ouôš* “to want” ⲧⲱⲙⲛⲧ *tômnt* “to meet”
ϭι-ⲗⲁⲡⲥⲓ *ci-lapsi* “to bite”	ⲗⲱⲕⲥ *lôks* “to bite”
ϯ-ϣⲉ-ⲛ-ⲧⲁⲡ *ti-še-n-tap* “to pierce”	ⲕⲱⲛⲥ *kôns* “to pierce”
ϫⲉⲙ-ⲛⲟⲙϯ / ϯ-ⲛⲟⲙϯ *čem-nomti / ti-nomti* “to be comforted, to comfort”	ⲥⲟⲗⲥⲗ *solsl* “to be comforted, to comfort”
ϯ-ϩⲟ *ti-ho* “to implore”	ⲥⲟⲡⲥ *sops* “to implore”

Table 7.2 Bohairic LVCs vs. Sahidic synthetic counterparts in the Biblical corpus.

9 The tables specify the meaning of the counterpart verbs, whenever it deviates from the translation of the LVC in question.

10 Psalm 118:39.

While the expressions common to both dialects may have been inherited from earlier language stages or else could represent a parallel development, the constructions with synthetic equivalents obviously belong to a more recent stratum.

Table 7.3 exemplifies cases where both dialects use different LVCs for (at least, roughly) one and the same meaning.

Sahidic	Bohairic	Translation
ⲉⲣ-ⲛⲟϭ *er-noc*	ⲉⲣ-ϩⲟⲩⲓⲧ *er-houit*	“to be(come) big, first”
ⲧⲁⲱⲉ-ⲟⲉⲓⲱ *taše-oeiš*	ϩⲓ-ⲱⲓⲱ *hi-ôiš*	“to proclaim, to preach”
ⲉⲣ-ⲱⲁⲩ *er-šau* ⲧⲱⲧ ⲛ-ϩⲏⲧ *tôt nhêt*	ⲧ-ⲙⲁⲧ *ti-mati*	“to fit, to agree, to suit”

Table 7.3 Sahidic LVCs vs. Bohairic LVCs.

As these examples show, the core of the light-verb inventory consists of the verbs (ⲉ)ⲣ- (*e)r-* “to do; to be”, ⲧ- *ti-* “to give”, and ⲭⲓ- *či-* “to take”, with significantly fewer examples of other verbs, for example, ϭⲛ- *cn-* “to find”, ⲕⲁ- *ka-* “to put”, ⲫⲓ- *fi-* “to lift”, ⲱⲉⲡ- *šep-* “to receive”, ⲭⲉ-, ⲭⲓ- *če-*, *či-* “to say”, ϩⲓ- *hi-* “to throw”. This set looks similar to those attested in many other languages, such as (post-)classical Greek (Fendel 2023), modern European languages of the PARSEME sample, for example, Irish (Bloch-Trojnar 2010), to some extent Salishan Bella Coola (Mithun 1997), and Chinese (Zhu 1985, Feng 2016).[11] A considerable body of research explores the possibility of a shared semantic or grammatical function among light verbs as a lexicogrammatical category. On different accounts, light verbs do (Butt 2003, Seiss 2009) or do not (Sobolak 2023) assign theta-roles. At least, in some languages, they create the effect of boundedness, ascribing telicity to the event in question (Sobolak 2023). Finally, they are semantically selective, being not totally devoid of the semantics of their own (Cetnarowska 1993, Butt 2003, Sobolak 2023). In English (Cetnarowska 1993) and in Irish (Bloch-Trojnar 2010), the choice of the light verb is shown to depend, among other things, on the theta-role orientation of the predicated event (i.e., whether the subject referent is construed as an agent, a patient, or an experiencer of the event) and, accordingly, on the encoded meaning of (un)intentionality. Interestingly, construing the subject as a patient or experiencer of the complex predicate would still be compatible with the telic aspect of the predicate: telicity, in this case, results from the transitivity of the light verb itself.

In Coptic, synthetic predicates and synonymous LVCs do not display aspectual preferences. This must stem from the fact that the Coptic verbal system does

11 See also Butt 2003: 18, Hook 1991: 60.

not support any aspectual ambiguity; aspects are encoded in the conjugational morphology. On the other hand, there is some correlation between a light verb and the theta-role of the subject, at least, in as much as *give*-compounds usually introduce a causative event, and *take*-compounds do not. More precisely, the three main light verbs differ in how much control or volitionality the subject referent has over the resulting event. *Give*-compounds are +controlled/volitional, *take*-compounds, as a rule, are controlled/volitional, *do*-compounds are neutral in this respect.

Though in a very few tokens, Coptic also manifests a phenomenon that is, to my knowledge, unattested in other languages, that is, secondary affixation, when a light verb is attached to a stem that already consists of a light verb and a noun.[12] The Bohairic Bible contains examples, such as ⲉⲣ-ϭⲓ-ⲭⲉⲛⲧ *er-či-fent* "to be worm-eaten" (Acts 12:23), ⲉⲣ-ϭⲓ-ⲛⲓⲟⲟⲣ *er-či-nior* "to ford a river" (Gen.31:21 and passim), and ⲉⲣ-ϭⲓ-ⲥⲁⲛⲓⲥ *er-či-sanis* "to have doubts" (Psalms 118:39), while in a fragment of the Sahidic *Shepherd of Hermas*, the compound ⲣ-ⲡ-ⲙⲉⲉⲩⲉ *er-p-meeue* "to remember" is nominalised by the definite article ⲡ- *p-* and further expanded by another ⲉⲓⲣⲉ *eire* "to do" (ⲉⲓⲣⲉ ⲙ-ⲡ-ⲣ-ⲡ-ⲙⲉⲉⲩⲉ ⲙ-ⲡⲉϥ-ⲙⲟⲩ *eire m-p-r-p-meeue m-pef-mou* "to remember his death" (Lefort 1940)).[13] The diachronic development of such LVCs can be reconstructed as follows: Coptic had a productive zero-derivation from verbs to nouns, so the verbalised expressions ϭⲓ-ⲭⲉⲛⲧ *či-fent* "to take worms" and ϭⲓ-ⲥⲁⲛⲓⲥ *či-sanis* "to take doubt" became univerbated and in their turn nominalised, which prompted the second-order verbal derivation. To appreciate this phenomenon in the larger context of Coptic morphology, we have to remember the specifics of Coptic word-formation. Coptic can simultaneously be termed an analytic (Funk 1978, Reintges 2011) and a synthetic (Haspelmath 2015) language due to the dual nature of its morphology. On the level of lexical units, Coptic is a root-and-pattern language. It almost completely lacks concatenative inflectional morphology and has a very restricted derivational morphology (Jernstedt 1986: 169-176, Richter 2023: 13). On the level of prosodic units, Coptic is an agglutinative language; its verb-containing unit can consist of two to five morphs. Light-verb doubling points to the fact that, at least, the primary light verbs in these lexemes came all the way from being a part of a syntactic construction (verbal phrase) to being an affix capable of further morphological integration. Thus, similarly to the abstract-deriving prefix ⲙⲛⲧ- *mnt-* and the privative prefix ⲁⲧ- *at-*, it can signal a transition from the root-and-pattern morphology inherited from earlier Egyptian to a concatenative

12 The phenomenon differs from light-verb stacking as discussed in Butt (1995). In Urdu, each additional light verb contributes a distinct semantic or grammatical meaning (e.g. a permissive operator scoping over an aspectual one; Butt 1995: 129), whereas in Coptic it functions solely as a derivational device.

13 The author is grateful to I. Miroshnikov for pointing out this example.

morphology. The precise stages of this grammaticalisation process, however, remain difficult to determine.

Occasionally, light verbs undergo phonetic changes, as seen in the construct form ϫⲉ- *če-* of the verb ϫⲱ *čô* "to say" in LVCs like ϫⲓ-ⲟⲩⲁ *či-oua* "to blaspheme" and ϫⲓ-ⲥⲟⲗ *či-col* "to lie".[14] This phonetic fusion provides further support for analysing light verbs as affixes rather than as independent constituents of the verb phrase (VP).

As for the nominal parts of Coptic light-verb compounds, they are not restricted to eventive nouns, such as ϫⲣⲟⲡ *črop* "hindrance", ⲙⲧⲟⲛ *mton* "rest", ⲍⲟⲧⲉ *hote* "fear", ⲯⲟϫⲛⲉ *šočne* "counsel", but include also the names of various body parts (ⲍⲏⲧ *hêt* "heart", ⲣⲟ *ro* "mouth", ⲣⲁⲧ= *rat* "foot", ⲧⲟⲟⲧ= *toot-* "hand") and of objects, such as ⲙⲟⲉⲓⲧ *moeit* "way" in ϫⲓ-ⲙⲟⲉⲓⲧ *či-moeit* "to lead" or ⲓⲟⲟⲣ *ioor* (from ⲓⲁⲣⲟ *iaro* "river" or ⲉⲓⲟⲟⲣ *eioor* "canal") in ϫⲓⲟⲟⲣ *čioor* "to ford (a river)". Interestingly, some nouns participating in LVCs are not attested outside of these constructions. Thus, ⲁϩⲟⲙ *ahom* exists only in ⲁⲯ-ⲁϩⲟⲙ *aš-ahom* (Sah) / ⲫⲓ-ⲁϩⲟⲙ *fi-ahom* (Bohairic) "to utter a sigh", ⲟⲉⲓⲯ / ⲱⲓⲯ *oeiš* / *ôiš* only in ⲧⲁⲯⲉ-ⲟⲉⲓⲯ *taše- ôiš* (Sahidic) / ϩⲓ-ⲱⲓⲯ *hi- ôiš* (Bohairic) "to cry out", and ϫⲛⲓⲧ *čnit*, only in ϫⲓ-ϫⲛⲓⲧ *či-čnit* "to try". Possibly, the nouns became obsolete after the process of univerbation was completed. The newer lexeme then could be nominalised, just as we saw happen within the double-prefix combinations, (12).

(12) Sahidic, Psalms 6:7

ⲁⲓϩⲓⲥⲉ			ϩⲙ	ⲡⲁⲁⲯⲁϩⲟⲙ	
a-	*i-*	*hise*	*hm*	*pa-*	*aš-ahom*
PST	1SG	suffer	with	POSS.SGM:1SG-	sigh

"I became weary with my moaning."

4. Valency of light-verb constructions

As distinct lexical units, light-verb compounds establish a valency pattern of their own. However, the semi-lexicalised nature of these constructions raises the question of which component ultimately determines the valency of the whole compound. Depending on the language, all four logically possible answers can be found in the literature. For Japanese, it is argued that a light verb is just a licensing host for the theta-roles assigned by the semantic head of the compound (Grimshaw and Mester 1988). Elsewhere, different light verbs can assign distinct valencies with one and the same semantic head thereby affecting

14 See also Miroshnikov (forthcoming) for Fayyumic: "Since the spelling ⲉⲣ- [*er-*] consistently occurs in combination with ϩⲙⲙⲓ [*hēmmi*], one might wonder whether the former element was no longer recognized as the status constructus of ⲓⲗⲓ [*ili*] but rather was seen as part of the stem."

the theta-role configuration of the construction. Seiss (2009) illustrates this by an example from Bardi (Nyulnyulan family) where the light verbs *ma* "to put" and *ga* "to take" derive verbs with different theta-role repertories from the stem *abarrabarr* "confusion":

1.	*abarrabarr-ma-* "to be careless"	Theme
2.	*abarrabarr-ga-* "to lead someone astray"	Agent, Patient

Kettnerová and Bejček (2016) claims that in Czech, both the light verb and the object noun contribute to the valency pattern of the compound.

Finally, in a study dedicated specifically to LVCs with lexemes of Greek origin in Bohairic Coptic, Zakrzewska claims that the use of prepositions other than the default marker ⲛ/ⲙⲙⲟ= *n-/mmo=* with the second argument "is limited to certain semantic classes of verbs, such as verbs of perception and cognition and verbs denoting contact-by-motion, whose second argument is marked by the allative preposition [ⲉ-/ⲉⲣⲟ=] *e-/ero=* or verbs of ritual behavior and speech actions, which occur with [ⲛ-/ⲛⲁ=] *n-/na=*" (Zakrzewska 2017: 233). In Zakrzewska's opinion, therefore, theta-roles are not assigned by the light verb or its complement, but stem from the semantic class of the resulting verbal lexeme.

As can be seen already from Zakrzewska's description, Coptic LVCs are not homogeneous with respect to their theta-role licensing. Rather, they form classes depending on the degree of univerbation. Some LVCs are fully or almost fully analysable and more construction-, than word-like. The valency in such compounds can be seen as determined by one element, less often the noun (as in ϭⲛ-ϩⲙⲟⲧ ⲙⲡⲉⲙⲧⲟ ⲛ- *cn-hmot mpemto n-* "to find grace before"), more often the verb, as in various compounds with ⲧ *t^{i}* "to give" (for example, ⲧ-ⲥⲃⲱ *t^{i}-sbô* "to give knowledge" = "to teach").

Quite a number of compounds with ⲧ *t^{i}* "to give" and ⲭⲓ *či* "to take" with one and the same complement noun occupying the position of the direct object form diathetic pairs where the oblique argument of *give*-compounds/the subject of *take*-compounds represents patient or beneficiary, or experiencer, while the subject of *give*-compounds refers to agent or stimulus. This class includes, for instance, ⲧ-ⲃⲉⲕⲉ / ⲭⲓ-ⲃⲉⲕⲉ *t^{i}-beke* / *či-beke* "to pay; to be paid", ⲧ-ⲙⲕⲁϩ / ϭⲓ-ⲙⲕⲁϩ *t^{i}-mkah* / *či-mkah* "to cause grief; to be grieved", ⲧ-ⲙⲧⲟⲛ / ϭⲓ-ⲙⲧⲟⲛ *t^{i}-mton /či-mton* "to give a rest; to have a rest", ⲧ-ⲡⲁϩⲣⲉ / ⲭⲓ-ⲡⲁϩⲣⲉ *t^{i}-pahre* / *či-pahre* "to cure; to be cured", ⲧ-ⲥⲃⲱ / ⲭⲓ-ⲥⲃⲱ *t^{i}-sbô* /*či-sbô* "to teach; to be taught", ⲧ-ⲥⲱϣ / ⲭⲓ-ⲥⲱϣ *t^{i}-sôš / či-sôš* "to despise; to be despised", ⲧ-ϣⲓⲡⲉ / ⲭⲓ-ϣⲓⲡⲉ *t^{i}-šipe /či-šipe* "to give shame; to be ashamed", ⲧ-ϧⲣⲟⲡ / ⲭⲓ-ϧⲣⲟⲡ *t^{i}-črop / či-črop* "to hinder; to be impeded". Greek-origin nouns can also serve as the basis for these derivations, for example, Sahidic ⲭⲓ-ⲃⲁⲡⲧⲓⲥⲙⲁ *či-baptisma* "to be baptised" vs. ⲧ-ⲃⲁⲡⲧⲓⲥⲙⲁ *t^{i}-baptisma* "to baptise" (cf. Bohairic ϭⲓ-ⲱⲙⲥ / ⲧ-ⲱⲙⲥ c*i-ômѕ* vs. *t^{i}-ômѕ*). In these

compounds, both light verbs preserve their initial argument structure (including the complement noun, *give* has three arguments and *take* has two), which makes the *give*-compounds bivalent, and the *take*-compounds intransitive. Typically, in such cases, ⲧ *t*^{i} "to give" assigns dative case (marked by the preposition ⲛ-/ ⲛⲁ= *n-/na=* in Coptic) to the second argument.

However, the parallel derivation of ⲧ-/ϫⲓ- *t*^{i}*- / či-* as inducing thematic roles of stimulus/(in)direct causer and experiencer/patient, respectively, is not regular or pervasive, thus does not reach the extent of a full-fledged grammatical device. So, the subjects of ⲧ-ⲟⲥⲉ *ti-ose* "to lose" (literally "to give loss"), ⲧ-ⲟⲩⲃⲁϣ *ti-oubaš* "to grow white" (literally "to give white") do not denote either a stimulus, or a causer, while the subjects of ϫⲓ-ϫⲛⲓⲧ *či-čnit* "to test, to examine" (literally "to take test") or ϫⲓ-ⲙⲟⲉⲓⲧ *či-moeit* "to lead" (literally "to take way") do not refer to patients.

What happens in those less easily analysable cases is a certain metaphorical reinterpretation of the compound which can be more or less remote from the sum of the literal meanings of the parts, as, for instance, in ϫⲓ-ⲥⲙⲏ *či-smê* "to hear" (literally "to take voice"), ϫⲓ-ϩⲟ *či-ho* "to favour" (literally "to take face"), ⲙⲉϩ-ⲣⲱ= *meh-rô=* "to bite" (literally "to fill mouth"), ϫⲓ-ⲏⲡⲉ *či-êpe* "to count" (literally "to take number"), ϫⲓ-ϣⲟϫⲛⲉ *či-šočne* "to advise" (literally "to take decision"). In cases like these, the argument structure is defined neither by the light verb, nor by the complement noun. Rather, it mirrors the argument structure of synthetic verbs with similar meanings and thus, in line with Zakrzewska's observations, is determined by the semantic class of the verb. The semantic gradations are subtle, so I shall limit myself to pointing out only a few regularities or particularly curious cases. The addressee of a speech act is typically coded with the dative ⲛ-/ⲛⲁ= *n- / na=* preposition, so this is the valency of ϫⲓ-ϣⲟϫⲛⲉ *či-šočne* "to advise", ϫⲓ-ⲡ-ⲟⲩⲱ *či-p-ouô* "to inform", (13).

(13) Sahidic, Acts 27:9

ⲡⲁⲩⲗⲟⲥ	ⲇⲉ	ⲁⲫϫⲓϣⲟϫⲛⲉ	ⲛⲁⲩ
paulos	*de*	*a-f-či-šočne*	*na=u*
Paulos	PTCL	PST-3SGM-take-decision	DAT-3PL

"And Paul advised them"

Neither ϣⲟϫⲛⲉ *šočne* "decision", nor ϫⲓ *či* "to take" license the theta-role of recipient, so the dative valency can only be explained as licensed by the compound as a whole.

The simplex verbs of perception take the allative ⲉ- / ⲉⲣⲟ= *e- / ero=* preposition, and so does the complex verb ϫⲓⲥⲙⲏ *či-smê* "to listen" (literally "to take voice"). Here, too, neither of the parts licenses the valency of the whole, (14).

(14) Sahidic, Genesis 4:23

ⲥⲱⲧⲙ	ⲉⲡⲁϩⲣⲟⲟⲩ	ⲛⲉϩⲓⲟⲙⲉ	ⲛⲗⲁⲙⲉⲭ
sôtm	*e-pa-hroou*	*ne-hiome*	*n-lamekh*
hear	ALL-POSS.SGM:1SG-voice	DEF.PL-women	GEN-Lamech

ϫⲓⲥⲙⲏ	ⲉⲛⲁⲯⲁϫⲉ
či-smê	*e-na-šače*
take-voice	ALL-POSS.PL:1SG-word

"Hear my voice, Lamech's wives, listen to my words."

ⲉ-/ ⲉⲣⲟ= *e-/ero=* is also the preferred preposition for the second argument, if the first argument is affected, as in ϫⲓ-ⲃⲟⲧⲉ *či-bote* "to feel aversion", ϯ-ⲥⲟ *t^{i}-so* "to feel pity" or ⲕⲁ-ϩⲧⲏ= *ka-htê=* "to lean, to confide", ⲥⲛ-ⲁⲣⲓⲕⲉ *cn-arike* "to reproach" (literally "to find fault"), (15).

(15) Sahidic, Psalms 15:1

ϩⲁⲣⲉϩ	ⲉⲣⲟⲓ	ⲡϫⲟⲉⲓⲥ	ϫⲉ	ⲁⲓⲕⲁϩⲧⲏⲓ
hareh	*ero=i*	*p-čoeis*	*če*	*a-i-ka-htê=i*
protect ALL-1SG	ALL-1SG	DEF.SGM-lord	because	PST-1SG-put-heart-1SG

ⲉⲣⲟⲕ
ero=k
ALL-2SGM

"Protect me, Lord, for I confided in you."

The occurrence of the direct object marking ⲛ-/ⲙⲙⲟ= *n-/mmo=* after the compound signals that the light verb and the complement noun have fully merged into a single lexical unit. This argument pattern is manifested with numerous LVCs, such as ϫⲓ-ⲕⲃⲁ *či-kba* "to take revenge", ϫⲓ-ⲙⲁⲉⲓⲛ *či-maein* "to make divination", ϫⲓⲟⲟⲣ *čioor* "to ford river", ϫⲓ-ⲕⲣⲟϥ *či-krof* and ⲣ-ϩⲁⲗ *r-hal* "to deceive", ϯ-ⲟⲥⲉ *ti-ose* "to lose", ϫⲓ-ⲁⲛⲓⲧ *či-čnit* "to test", ⲥⲙⲛ-ⲥⲛⲧⲉ *smn-snte* "to establish", ⲣ-ⲭⲣⲉⲓⲁ *r-xria* "to need", ϫⲓ-ϩⲱⲛ *či-hôn* "to betroth", ⲣ-ⲯⲡⲏⲣⲉ *r-špere* "to amaze", ⲟⲩⲉϩ-ⲥⲁϩⲛⲉ *oueh-sahne* "to order" etc., (16) and (17).

(16) Sahidic, Leviticus 26:25

ϯⲛⲁⲉⲓⲛⲉ	ⲛⲟⲩⲥⲏⲫⲉ	ⲉϫⲛ ⲑⲩⲧⲛ	ⲉϫⲓⲕⲃⲁ
ti-na-eine	*n-ou-sêfe*	*ečn-têuten*	*e-či-kba*
1SG-FUT-bring	ACC-SG.INDEF-sword	upon-2PL	to-take-vengeance

ⲙⲙⲱⲧⲛ	ⲛⲟⲩϩⲁⲡ	ⲛⲧⲁⲇⲓⲁⲑⲏⲕⲏ
mmô=tn	*n-ou-hap*	*n-ta-diathêkê*
ACC-2PL	ACC-SG.INDEF-judgement	GEN-POSS.SGF:1SG-covenant

"And I will bring the sword on you to take vengeance on you for judging my covenant."

(17) Sahidic, Hosea 2:21

ϯⲛⲁⲭⲓϩⲱⲛ	ⲙⲙⲟ	ⲛⲁⲓ	ⲱⲁⲉⲛⲉϩ
ti-na-či-hôn	*mmo=Ø*	*na=i*	*ša-eneh*
1SG-FUT-take nuptial	ACC-2SGF	DAT-1SG	till-eternity

"I will betroth you to me for eternity."

With some LVCs, a pronominal P[atient]-argument can be indexed as the possessor in the possessive prefix of the complement noun[15]; ϥⲓ-ⲡⲉϥ-ⲣⲟⲟⲩⲱ *fi-pef-roouš* (literally "to carry his care") means "to take care of him", ⲣ-ⲡⲉϥ-ⲙⲉⲉⲩⲉ *r-pef-meeue* (literally "to do his thought") means "to remember him", and so on, (18) and (19).

(18) Sahidic, Deuteronomy 20:5

ⲛⲧⲉⲕⲉⲣⲱⲙⲉ	ⲭⲓⲡⲉϥⲁⲉⲓⲕ	ⲉⲃⲟⲗ
nte-ke-rome	*či-pe=f-aeik*	*ebol*
CONJ-other-man	take-POSS.SGM:3SGM-consecration	out

"... and someone else starts inhabiting it."

(19) Sahidic, Numbers 1:49

ⲁⲛⲁⲩ	ⲉⲧⲉⲫⲩⲗⲏ	ⲛⲗⲉⲩⲉⲓ	ⲛⲛⲉⲕⲥⲙⲡⲉⲥⲱⲓⲛⲉ
anau	*e-te-phulê*	*n-leuei*	*nne=k-cm-pe=s-šine*
see	ALL-DEF.SGF-tribe	of-Levi	NEG.OPT-2SGM-find-POSS.SGM:3SGF-seeking

ⲁⲩⲱ	ⲛⲛⲉⲕⲭⲓⲧⲉⲩⲏⲡⲉ	ϩⲛ	ⲧⲙⲏⲧⲉ
auô	*nne=k-či-te=u-epe*	*hn*	*t-mête*
and	NEG.OPT-2SGM-take-POSS.SGF:3PL-number	in	DEF.SGF-middle

ⲛⲛⲱⲏⲣⲉ	ⲙⲡⲓⲥⲣⲁⲏⲗ
n-n-šêre	*m–p-israêl*
GEN-DEF.PL-son	GEN-DEF.SGM-Israel

"Watch out for the tribe of Levi, do not visit it and do not count it among the children of Israel."

In the biblical corpus, I found ten such LVCs (besides the ones already mentioned, these are ⲥⲙ-ⲡ-ⲱⲓⲛⲉ *cm-p=šine* "to visit", ⲣ-ⲡ-ⲱⲃⲱ *r-p=ôbš* "to forget", ⲣ-ⲡ-ϩⲛⲃⲉ *r-p=hêbe* "to mourn", ⲣ-ⲡ-ϩⲁⲡ/ ϥⲓ-ⲡ-ϩⲁⲡ *r-p=hap / fi-p=hap* "to judge", ⲭⲓ-ⲡ-ⲁⲉⲓⲕ *či-p=aeik* "to consecrate, inhabit", ⲭⲓ-ⲡ-ⲕⲃⲁ *či-p=kba* "to avenge", ⲭⲓ-ⲡ-ⲛⲡⲉ *či-p=êpe* "to count", and ⲣ-ⲡ-ⲃⲟⲗ *r-p=bol* "to leave"). Two of them (ⲭⲓ-ⲡ=ⲕⲃⲁ *či-p=kba* "to avenge" and ⲭⲓ-ⲡ-ⲛⲡⲉ *či-p=êpe* "to count") alternate with the canonical transitive pattern, as ⲭⲓ-ⲕⲃⲁ ⲙⲙⲱⲧⲛ *ci-kba mmôtn* in (16) above vs. ⲭⲓ ⲙ-ⲡⲉϥ-ⲕⲃⲁ *ci m-pef-kba* in (20).

15 The same phenomenon is attested in Kurdish (Haig 2002) and Mano (Khachaturyan 2017).

(20) Sahidic, Exodus 21.20

ϩⲛ	ⲟⲩϫⲓⲕⲃⲁ	ⲉⲩⲉϫⲓ	ⲙⲡⲉϥⲕⲃⲁ
hn	*ou-či-kba*	*e=u-e-či*	*m-pef-kba*
with	SG.INDEF-take-revenge	OPT_3PL-take	ACC-POSS.SGM:3SGM-revenge

"(When a man strikes his slave, male or female, with a rod and the slave dies under his hand,) let him be punished with a revenge."

Nominal arguments attach via the preposition ⲛ- *n-* in both transitive subpatterns.

Finally, with LVCs whose inner complement is an inalienable noun, the non-subject argument can be indexed as the suffix on the complement noun. Some LVCs supporting this pattern are ⲧ-ⲣⲓⲛ= *t^{i}-rin*= "to name", ⲧⲁⲩⲉ-ⲣⲓⲛ= *taue-rin* "to pronounce name", ϫⲓ-ⲣⲱ= *či-rô*= "to silence; to capture" (literally "to take mouth"), ϣⲉⲡ-ⲧⲟⲟⲧ= *šep-toot*= "to betroth" (literally "receive hand"), ⲧⲟⲩⲛ-ⲉⲓⲁⲧ= *toun-eiat*= "to inform, to teach" (literally "to raise eyes"), ⲣ-ⲁⲛⲁ= *r-ana*= "to please". Despite becoming formally definite, the suffixed noun does not become referential, i.e., does not denote any specific entity. For instance, ⲧⲟⲩⲛⲉⲓⲁⲧϥ *toun-eiat=f* does not involve the reference to Joseph's eyes in (21).

(21) Sahidic, Matthew 2:22

ⲁⲩⲧⲟⲩⲛⲉⲓⲁⲧϥ			ⲇⲉ	ⲉⲃⲟⲗ	ϩⲛ	ⲟⲩⲣⲁⲥⲟⲩ
a=	*u-*	*toun-eiat=f*	*de*	*ebol*	*hn*	*ou-rasou*
PST-	3PL-	raise-eye-3SGM.POSS	PTCL	out	in	SG.INDEF-dream

"He was informed in his dream…"

The syntactic position of the suffix in these LVCs is not reserved for the direct object: the suffix can refer back to the subject, as well. The distinction is purely lexical. Thus, in ⲙⲟⲩϩ-ⲉⲓⲁⲧ= *mouh-eiat*= "to look intently" (literally "to fill eye") the suffix is co-referent with the subject, whereas in ⲧⲟⲩⲛ-ⲉⲓⲁⲧ= *toun-eiat*= "to inform, to teach" (literally "to raise eye"), it refers to the object. The same distinction holds between ⲕⲁ-ⲣⲱ= *ka-rô*= "to become silent" (literally "to put mouth") and ϫⲓ-ⲣⲱ= *či-rô*= "to make silent" (literally "to take mouth"), respectively. A particularly striking example is the LVC ⲧ-ⲧⲟⲟⲧ= *t^{i}-toot*= (literally "to give hand"), which exhibits a sharp grammatical and semantic distinction depending on the prepositional marking. In ⲧ-ⲧⲟⲟⲧ= ⲙⲛ- *t^{i}-toot*= *mn-* ("to give hand with"), the suffix refers anaphorically to the subject, and the construction means "to help", (22). In contrast, with ⲧ-ⲧⲟⲟⲧ= *t^{i}-toot*= followed by ⲛ- *n-*, ⲉ- *e-*, or ⲉⲣⲛ- *ern-*, the meaning shifts to "to get hold of something" or "to grasp". Here, too, the suffix is co-referential with the subject, (23). Finally, the unexpanded

form ϯ-ⲧⲟⲟⲧ= *t^{i}-toot=* also conveys the meaning “to help”, but in this case the suffix marks the P[atient]-argument, (24).

(22) Sahidic, Deuteronomy 1.30

ⲛⲧⲟϥ	ⲡⲉⲧⲛⲁϯⲧⲟⲟⲧϥ	ⲛⲙⲙⲏⲧⲛ	ⲉⲙⲓⲯⲉ	ⲛⲙⲙⲁⲩ
ntof	*p-et-na-ti-toot=f*	*nmmê=tn*	*e-miše*	*nemma=u*
he	DEF.SGM-REL-FUT-give-hand-3SGM.POSS	with-2PL	to-fight	with-3PL

“He will help you fight them.”

(23) Sahidic, Deuteronomy 9.17

ⲁⲓϯⲧⲟⲟⲧ	ⲛⲧⲉⲡⲗⲁⲝ	ⲥⲛⲧⲉ	ⲛⲱⲛⲉ
a-i-ti-toot	*n-te-plaks*	*snte*	*n-ône*
PST-1SG-give-hand.1SG	ACC-DEF.SGF-tablet	two.F	ATTR-stone

“I took hold of the two stone tablets.”

(24) Sahidic, Romans 8.28

ⲛⲉⲧⲙⲉ			ⲙⲡⲛⲟⲩⲧⲉ		ⲯⲁⲣⲉⲡⲛⲟⲩⲧⲉ
n-	*et-*	*me*	*m-*	*p-noute*	*šare-p-noute*
DEF.PL-	REL-	love	ACC-	DEF.SGM-god	HAB-DEF.SGM-god

ϯⲧⲟⲟⲧⲟⲩ	ϩⲛ	ϩⲱⲃ	ⲛⲓⲙ	ⲉⲡⲁⲅⲁⲑⲟⲛ	
ti-toot=ou	*hn*	*hôb*	*nim*	*e-*	*p-agathon*
give-hand-3PL	in	thing	every	ALL-	DEF.SGM-good

“Whoever loves God, God helps them to everything good.”

As we see, the obligatory inalienability suffix may be interpreted as a second argument, but only when the compound’s meaning licenses such reference. In other words, the noun’s inalienability alone does not render the construction transitive; valency is determined by the compound as a whole.

Unlike the quasi-LVCs ϫⲓ-ⲛ-ⲥⲟⲛⲥ *či-n-cons* and ϫⲓ ⲛ-ϫⲛⲁϩ *či-n-čnah* “to make violence”,[16] for which all the forms of the light verb (both construct forms and stative) are attested, true Coptic LVCs exhibit almost no morphological alternation in the verb. This is cross-linguistically unusual and places them closer to lexicalisations on the scale between LVCs and lexicalisations. A notable exception is the use of morphological imperatives ⲁⲣⲓ- *ari-* for ⲣ-/ⲉⲣ- *r-*/*er-* “to do” in both

16 The two constructions mentioned constitute a subclass of their own. Unlike LVCs, they consist of a verb and an adverbial modifier, which may explain their freer morphosyntax (they can, e.g., take an immediate (pro)nominal object) and their lesser degree of lexicalisation. They are not discussed here for reasons of space.

Sahidic and Bohairic and ⲙⲁ- *ma-* for ϯ- *tⁱ-* “to give” in Bohairic, (25) and (26).

(25) Bohairic, Job 6:24

ⲙⲁⲥⲃⲱ	ⲛⲏⲓ	ⲁⲛⲟⲕ	ⲇⲉ	ϯⲛⲁⲭⲁⲣⲱⲓ
ma-sbô	*nê=i*	*anok*	*de*	*ti-na-kha-rô=i*
give.IMP- knowledge	DAT-1SG	I	PTCL	1SG-FUT-put-mouth-1SG

“Teach me, and I will be quiet.”

(26) Bohairic, Philemon 1:20

ⲙⲁⲙⲧⲟⲛ	ⲛⲛⲁⲙⲉⲧⲯⲁⲛⲑⲙⲁϩⲧ	ϧⲉⲛ	ⲡⲭ̅ⲥ̅
ma-mton	*n-na-metšanthmaxt*	*xen*	*p-kh(risto)s*
give.IMP-rest	DAT-POSS.PL:1SG-compassion	in	DEF.SGM-Christ

“Refresh my compassion in Christ.”

In Sahidic, the morphological infinitive ϯ *ti* “to give” is used in these cases, as opposed to the full verb that uses the imperative ⲙⲁ *ma*.

Univerbation of LVCs sometimes prompts the analogical filling of what is perceived as gaps in their morphological paradigms. Thus, ⲕⲁ-ⲣⲱ= *ka-rô=* “to keep silence” (literally “to put mouth”) acquires the stative form ⲕⲁⲣⲁⲉⲓⲧ *karaeit* built by analogy to such stative forms as ϩⲕⲁⲉⲓⲧ *hkaeit* from ϩⲕⲟ *hko* “to be hungry”, (27).

(27) Sahidic, Genesis 24.21

ⲡⲣⲱⲙⲉ	ⲇⲉ	ⲛⲉⲫⲙⲟⲩⲯⲧ	ⲙⲙⲟⲥ
p-rôme	*de*	*ne=f-moušt*	*mmo=s*
DEF.SGM-man	PTCL	PRET-3SGM-watch	ACC-3SGF

ⲉⲫⲕⲁⲣⲁⲉⲓⲧ
e=f-karaeit
CIRC-3SGM-silent

“The man was watching her silently.”

Another such instance is quoted in (Crum 1939: 648a): for ⲭⲓ-ϩⲣⲁ *či-hra=* “to amuse oneself” (literally “to take face”), there is some evidence for the existence of a stative ⲭⲓϩⲣⲁⲉⲓⲧ *čihraeit* “to be amused”. The scarcity of these examples, however, suggests that such morphological innovations were only marginal. Alternatively, when passivisation was required, the complement noun could be promoted to subject position, with the underlying P[atient]-argument indexed on it via a suffix, and the verb appearing in its stative form, (29) as opposed to (28)—as if the LVC were a regular verb-direct object phrase

and the complement noun were referential.

(28) Sahidic, Deuteronomy 22:23

ⲟⲩϣⲉⲉⲣⲉ	ϣⲏⲙ	ⲙⲡⲁⲣⲑⲉⲛⲟⲥ	ⲉⲁⲩϣⲡⲧⲟⲟⲧⲥ
ou-šeere	*šêm*	*m-parthenos*	*e-a=u-šp-toot=s*
SG.INDEF-girl	little	ATTR-virgin	CIRC-PST-3PL-receive-hand-3SGF

ⲛϩⲁⲓ
n-hai
DAT-husband

“a virgin girl who is betrothed…”

(29) Sahidic, Luke 2:5

ⲙⲁⲣⲓⲁ	ⲧⲉⲧⲉⲣⲉⲧⲟⲟⲧⲥ	ϣⲏⲡ	ⲛⲁϥ
maria	*t-etere-toot=s*	*šêp*	*na=f*
Maria	DEF.SGF-REL-hand-3SGF	receive.STAT	DAT-3SGM

“Maria who was betrothed to him…”

Since such instances are extremely infrequent, too, we must assume that diathetic alternation is not fully available even for transitive LVCs.

5. Light-verb constructions vs. synthetic verbs

Although the existence of a synthetic verb with a similar or equivalent meaning is often proposed as one of the criteria for identifying LVCs, to my knowledge, relatively few studies have systematically compared synthetic and analytic counterparts. Sobolak (2023) undertakes such a comparison for English and Korean, concluding that LVCs re-organise the event structure making it telic. Cetnarowska (1993: 41–42) assumes that the choice of an LVC in English may be influenced by discourse principles of end-weight and end-focus (as suggested in Quirk et al. 1972) or by the relative ease of modification (as claimed by Jespersen 1954 and Halliday 1985).

In Coptic, relatively few LVCs have a synthetic equivalent. In most cases, LVCs introduce entirely new concepts by verbalising nouns—for example, ⲣ-ⲭⲣⲉⲓⲁ *r-k^{h}reia* (“to need”) or ⲭⲓ-ⲙⲁⲧⲟⲩ *či-matou* (“to be poisoned”). When a close synthetic synonym does exist, in some cases, LVCs seem to be associated with a particular type of participants or situation. Thus, the verb ⲱⲡ *ôp* “to count” is consistently applied to inanimate/non-human entities, such as time units, stars, or cattle, (30) and (31).

(30) Sahidic, Deuteronomy 16:9

ⲧⲁⲓ	ⲧⲉ	ⲑⲉ	ⲉⲧⲕⲛⲁϣⲡ	ⲛⲥⲁϣϥⲉ	ⲛϩⲉⲃⲇⲟⲙⲁⲥ
tai	*te*	*t-he*	*et=k-na-ôp*	*n-sašfe*	*n-hebdomas*
this	COP.F	DEF.SGF-way	REL-2SGM-FUT-count	ACC-seven	GEN-week

“This is how you shall count seven weeks.”

(31) Sahidic, Isaiah 33:18

ⲉϥⲧⲱⲛ	ⲡⲉⲧⲱⲡ	ⲛⲛⲥⲓⲟⲩ
e=f-tôn	*p-et-ôp*	*n-n-siou*
FOC.PRS-3SGM-where	DEF.SGM-REL-count	ACC-DEF.PL-star

“Where is he who counts the stars?”

On the other hand, the object referents of the LVC ϫⲓ-ⲏⲡⲉ *či-êpe* “to count, to take number” are human, (32).

(32) Sahidic, Numbers 3:15

ϫⲓⲏⲡⲉ	ⲛⲛϣⲏⲣⲉ	ⲛⲗⲉⲩⲓ
či-êpe	*n-n-šêre*	*n-leui*
take-number	ACC-DEF.PL-son	GEN-levy

“Count the sons of Levi.”

This contrast, however, may not be solely attributable to animacy, for example, [+human], but to the type of action involved (ωπ *ôp* translates Greek ἀριθμέω *arithmeô* / Hebrew לספור *lispor* “to count”, while ϫⲓ-ⲏⲡⲉ *či-êpe* translates ἐπισκέπτομαι *episkeptomai* “to inspect” / לפקוד *lip^{h}kod* “to number”). Generally, it seems that in Coptic, LVCs serve as a productive means for lexical derivation, rather than vehicles for encoding consistent grammatical or structural features.

6. Coptic light-verb constructions and the Principle of Lexical Integrity

Let us consider again the LVCs in which the light verb and the complement noun can be separated by a possessive prefix referring to the P[atient]-argument, such as ⲣ-ⲡⲉϥ-ⲙⲉⲉⲩⲉ *r-pef-meeue* “to remember him” (literally “to make his thought”) or ⲥⲙ-ⲡⲉⲥ-ϣⲓⲛⲉ *cm-pes-šine* “to visit her” (literally “to find her seeking”). Understanding such constructions as minimal lexical units would contradict the Lexical Integrity Principle as interpreted in Booij (2009):[17]

17 Another instantiation of LVCs contradicting the Lexical Integrity Principle is discussed in Khachaturyan (2017: 9-10) with respect to Mano (Mande family). Here, the nominal parts

> "The principle of Lexical Integrity […] excludes two kinds of syntax-morphology interaction: (i) manipulation of parts of word-internal structure, and (ii) access to word-internal structure. Under manipulation I subsume the syntactic movement of a word constituent, and the *splitting up of words by intermediate constituents*." (my emphasis)

Similar infiltration of syntactic marking into morphology is not peculiar for Coptic LVCs; it is attested in much more word-like structures, *viz.*, in abstract nouns formed with the prefix ⲙⲛⲧ- *mnt-*. If the stem of the noun includes a morphologically changeable element, such nouns can mark the person/number, (33), or the gender, (34) and (35), of the possessor of quality:

(33a) Sahidic, Shenute, Canon 3 YB 330:21

ⲧⲉⲩⲙⲛⲧⲣⲉϥⲕⲁϩⲧⲏⲩ
te=u-mnt-ref-ka-hte=u
POSS.SGF:3PL-ABS-AGT-put-heart-3PL

"their confidence."

(33b) Sahidic, Leipoldt 1913:29:9

ⲧⲉⲩⲙⲛⲧⲙⲁⲓⲡⲉⲩⲙⲧⲟⲛ
te=u-mnt-mai-pe=u-mton
POSS.SGF:3PL-ABS-love.PART-POSS.M-3PL-rest

"their self-indulgence."

(34) Sahidic, Genesis 21:2

ⲁⲥⲭⲡⲟ	ⲛⲟⲩⲱⲏⲣⲉ	ⲛⲁⲃⲣⲁϩⲁⲙ	ϩⲛ
a=s-čpo	*n-ou-šêre*	*n-abraham*	*hn*
PST-3SGF-bear_child	ACC-SG.INDEF-boy	DAT-Abraham	in

ⲧⲉϥⲙⲛⲧϩⲗⲗⲟ
*tef-mnt-**hllo***
POSS.SGF:3SGM-ABS-old.M

"She bore a son to Abraham in his old age."

(35) Sahidic, Genesis 21:7

ⲁⲓⲭⲡⲉⲟⲩⲱⲏⲣⲉ	ϩⲛ	ⲧⲁⲙⲛⲧϩⲗⲗⲱ
a=i-čpe-ou-šêre	*hn*	*ta-mnt-**hllô***
PST-1SG-bear_child-SG.INDEF-boy	in	POSS.SGF:1SG-ABS-old.**F**

"I have born a son in my old age."

---

of LVCs can serve as referents for anaphoric pronouns. Obviously, both Coptic and Mano phenomena stem from the intermediate status of LVCs between words and constructions.

The prefix ⲙⲛⲧ- *mnt-* is a fully grammaticalised, bound derivational morpheme; it does not carry independent lexical meaning but functions as the morphological head of the derived lexeme assigning it feminine gender, so functionally it is equivalent to an abstract-deriving affix in European languages. Yet, we saw that, similarly to LVCs with an internal inflection, ⲙⲛⲧ- *mnt-*nominalisations remain accessible to agreement and indexing operations, revealing a mismatch between syntactic and morphological processes. Clearly, the lexical units, such as LVCs and ⲙⲛⲧ- *mnt-*nominalisations, occupy some middle field between constructions and words, pointing to the univerbation tendency which in time could change the morphological type of Coptic to concatenative.

7. Conclusion

Though the majority of native verbal vocabulary have bound transitive allomorphs, only a closed set of verbs can also function as light verbs, or in other words, verb-deriving affixes. Defined by *a priori* criteria, such as non-referentiality of noun objects, semantic voidness of the verb, and idiomaticity of the combination, some LVCs show additional properties testifying to the complete univerbation of the initial affix-noun combination: the ability to take arguments cognate to the object noun, the phonological assimilation of the verb, the generation of stative allomorphs to the verb-noun complex as a whole, and finally, occasional secondary derivation of an LVC from another LVC.

The core of the set of verb-deriving affixes consists of the verbs meaning *do*, *give*, and *take;* the last two frequently derive causative-passive pairs. The inventory of verb-deriving affixes has many cross-linguistic parallels. A promising line of inquiry would be to examine whether the lexical meanings of the verbs that become verbalisers are systematically associated with specific semantic effects in derivation—for instance, in Coptic, give tends to correlate with volitionality/agency, whereas take predominantly yields non-volitional/non-causative verb phrases.

As for the inventories of derivations, Sahidic and Bohairic differ: an LVC in one dialect often corresponds to a synthetic verb in the other. Yet in both dialects, this type of derivation represents a productive mechanism, applying to both native and borrowed nouns. Both examined dialects also exhibit the diachronic residue of this mechanism in the form of LVCs containing nouns that are not used outside these constructions. Synchronically, these function as semi-analysable lexicalised compounds.

Valency is rarely determined by the noun; more often it is imposed by the verb, and even more often, it matches the theta-roles defined by the meaning of the derived verb. In a few cases, the argument is morphosyntactically realised as a noun modifier (possessive prefix or suffix). The occurrence of an inflected

infix splitting a biradical verbal lexeme seems to contradict the Lexical Integrity Principle—a phenomenon that I attribute to the transitional character of Coptic morphology, which was probably developing in the direction of greater concatenativity.

Acknowledgements

The author wishes to express her gratitude to V. Fendel, L. Finkelberg, and I. Miroshnikov for their detailed linguistic comments and generous assistance. I am also deeply grateful to my anonymous reviewers whose insightful and meticulous comments have helped me to improve the paper.

Abbreviations

ABS abstractive
ACC accusative
ALL allative
ART article
CIRC circumstantial
CONJ conjunctive
COP copula
DAT dative
DEF definite
F feminine
FOC focus
FUT future
GEN genitive
HAB habitual
IMP imperative
INDEF indefinite
M masculine
NEG negative
OPT optative
POSS possessive
PRET preterite
PST past
PTCL particle
REL relative
SG singular
STAT stative

Bibliography

Bloch-Trojnar, Maria. 2010. Semantic Constraints on Light Verb Constructions in Modern Irish. *Studia Celto-Slavica* 4. 67–88. https://doi.org/10.54586/jbds2501

Booij, Geert. 2009. Lexical Integrity as a Formal Universal: A Constructionist View. In Sergio Scalise, Elisabetta Magni, and Antonietta Bisetto (eds.), *Universals of Language Today*, 83–100. Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-8825-4_5

Bresnan, Joan, and Sam Mchombo. 1995. The Lexical Integrity Principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13 (2). 181–254.

Butt, Miriam. 2003. The Light Verb Jungle. *Harvard Working Papers in Linguistics* 9. 1–49.

Cetnarowska, Bozena. 1993. *The Syntax, Semantics and Derivation of Bare Nominalisations in English*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Crum, Walter E. 1939. *Coptic Dictionary.* Oxford: Clarendon Press.

Fendel, Victoria Beatrix. 2022. The Missing Piece in the Jigsaw Puzzle: A Psycholinguistic Account of the Beginnings of the Coptic Alphabet. *Written Language and Literacy* 24(2). 198–228. https://doi.org/10.1075/wll.00053.fen

— 2023. Support-Verb Constructions with Objects: Greek-Coptic Interference in the Documentary Papyri? *Transactions of the Philological Society* 121(3). 382–403. https://doi.org/10.1111/1467-968x.12279

Feng, Shuang. 2016. *A Preliminary Theory of Diachronic Syntax in Chinese*. Shanghai: Shanghai Educational Publishing House.

Funk, Wolf-Peter. 1978. Toward a Synchronic Morphology of Coptic. In Robert McLachlan Wilson (ed.), *The Future of Coptic Studies*, 104–124. Coptic Studies 1. Leiden: Brill.

Grimshaw, Jane, and Armin Mester. 1988. Light Verbs and θ-Marking. *Linguistic Inquiry* 19 (2). 205–232.

Grossman, Eitan and Martin Haspelmath. 2014. The Leipzig-Jerusalem Transliteration of Coptic. In Eitan Grossman, Martin Haspelmath, and Tonia Sebastian Richter (eds.), *Egyptian-Coptic Linguistics in Typological Perspective*, 145–153. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110346510.145

Grossman, Eitan. 2023. Transitive Verbs as Lexical Affixes in Coptic. Paper presented at the Affixes Symposium. Turku, 18 August. https://sites.utu.fi/affixes/wp-content/uploads/sites/1247/2023/09/Grossman.pdf

Haig, Geoffrey. 2002. Noun-plus-verb Complex Predicates in Kurmanjî Kurdish: Argument Sharing, Argument Incorporation, or What? *STUF—Language Typology and Universals* 55 (1). 15–48.

Halliday, Michael A. K. 1985. *An Introduction to Functional Grammar*. London: Edward Arnold.

Haspelmath, Martin. 2015. A Grammatical Overview of Egyptian and Coptic. In Eitan Grossman, Martin Haspelmath and Tonio Sebastian Richter (eds.), *Egyptian-Coptic Linguistics in Typological Perspective*, 103–143. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110346510.103

— 2025. Compound and Incorporation Constructions as Combinations of Unexpandable Roots. *Zeitschrift für Wortbildung/Journal of Word Formation* 9(1). 1–23. https://doi.org/10.21248/zwjw.2025.1.124

Hook, Peter E. 1991. The Emergence of Perfective Aspect in Indo-Aryan. In Bernd Heine and Elizabeth Closs Traugott (eds.), *Approaches to Grammaticalization II: Types of Grammatical Markers*, 59–89. Amsterdam and Philadelphia, PN: John Benjamins.

Jernstedt, Petr V. 1986. *Issledovanija po grammatike koptskogo jazyka* (Studies in the Grammar of Coptic Language). Moscow: Nauka.

Jespersen, Otto. 1954. *A Modern English Grammar*. Vol. VI. London: George Allen and Unwin Ltd.

Kettnerová, Václava, and Eduard Bejček. 2016. Distribution of Valency Complements in Czech Complex Predicates: Between Verb and Noun. In Nicoletta Calzolari et al. (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16),* 515–521. Portorož: European Language Resource Association.

Khachaturyan, Maria. 2017. Bound Noun Plus Verb Combinations in Mano. *Mandenkan* 57. 3–24.

Layton, Bentley. 2011. *A Coptic Grammar: With Chrestomathy and Glossary; Sahidic Dialect.* 3rd rev. edn. Wiesbaden: Harrassowitz.Lefort, Louis Th. 1940. *Les manuscrits coptes de l'université de Louvain. Textes littéraires décrits par L. Th. Lefort.* Louvain: Bibliotheque de l'Université.Leipoldt, Iohannes. 1913. *Sinuthii archimandritae vita et opera omnia iv*. CSCO 73/Scr. Copt. 5. Louvain (reprinted 1954).

Miroshnikov, Ivan (forthcoming). *The Life of Onuphrius and the Martyrdom of Paphnutius in Fayyumic Coptic.*

Mithun, Marianne. 1984. The Evolution of Noun Incorporation. *Language* 60(4). 847–894.

— 1997. Lexical Affixes. In Andrew Spencer and Arnold M. Zwicky (eds.), *Handbook of Morphology*, 571–97. Oxford: Blackwell.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1972. *A Grammar of Contemporary English*. London: Longman.

Reintges, Chris H. 2011. High Analyticity and Coptic Particle Syntax: A Phase-Based Approach. *The Linguistic Review* 28(4). 533–599.

— 2018. *Coptic Egyptian (Sahidic Dialect): A Learner's Grammar*. Cologne: Köppe.

Richter, Tonio S. 2023. Coptic. In Andréas Stauder and Willeke Wendrich (eds.), *UCLA Encyclopedia of Egyptology*. Los Angeles, CA. https://escholarship.org/uc/item/qt22r6s881

Seiss, Melanie. 2009. On the Difference between Auxiliaries, Serial Verbs and Light Verbs. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG09 Conference*, 501–519. Stanford, CA: CSLI Publications.

Sobolak, Frances. 2023. The Event Structure of Light Verbs. PhD thesis (Cornell University).

Speransky, Nina. 2022. *Transitivity and Aspect in Sahidic Coptic. Studies in the Morphosyntax of Native and Greek-Origin Verbs.* Hamburg: Widmaier Verlag. https://doi.org/10.37011/studmon.26

Zakrzewska, Ewa. 2017. Complex Verbs in Bohairic Coptic: Language Contact and Valency. Brian Nolan and Elke Diedrichsen (eds.), *Argument Realisation in Complex Predicates and Complex Events: Verb-Verb Constructions at the Syntax-Semantic Interface*, 213–243. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.180.08zak

Zhu, Dexi. 1985. Xiandai Shumian Hanyu Li de Xuhua Dongci He Mingdongci (Dummy Verbs and Nominal Verbs in Modern Written Chinese). *Journal of Peking University (Humanities and Social Sciences)* 5. 1–6.

III. CRITICAL ENGAGEMENT: MORPHOSYNTACTIC, SEMANTIC, AND PRAGMATIC PERSPECTIVES

8. Roundabout ways of setting up scenes in the Greek New Testament (1st to 2nd c. CE)

Cressida Ryan[1]

1. Introduction

Luke opens his gospel with a prefatory declaration to Theophilus about his storytelling mission, which alerts his readers to a potential curation of and sensitivity to the ways in which Greek is able to tell stories. As we read the gospel, we are taken on a journey, from Nazareth to Jerusalem, and through Jesus' teaching. This journey, I suggest, is partly scaffolded by Luke's idiosyncratic use of the imperfect tense, and the multi-word (sometimes periphrastic) imperfect construction in particular. In this chapter, therefore, I compare and contrast the use of the multi-word imperfect as opposed to the synthetic imperfect in the gospels of Mark and Luke.[2] A fuller analysis including at least Acts of the Apostles and the other gospels would be interesting for the sake of contextualising the argument. Given that Acts deals with different kinds of materials, however, and John's gospel is very different to the synoptics, a comparison of Mark and Luke offers a useful and manageable starting point for the kind of analysis which can be undertaken in a short chapter.

The first section of this chapter outlines my methodology and approach. I define my understanding of the multi-word form of the imperfect, within an understanding of what does or does not class as "periphrastic", acknowledging where some disagreement lies. I then introduce my data and its limitations.

Before analysing the different constructions further, I then outline and explore the standard uses of the imperfect. I move on to use this to formulate a specific argument about how the multi-word version of the imperfect tense, however

1 The data underlying this chapter is available at: https://doi.org/10.5281/zenodo.18510208

2 See also du Toit (2017) on the imperfect in John's narrative discourse, as a starting point for other gospels.

©2026 Cressida Ryan, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.08

categorised, plays a specific narratological (and consequently exegetical) role in Mark and Luke, exploring key similarities and differences. Previous research has focussed primarily on understanding the imperfect in general, or on discerning what a periphrastic construction is, and for the New Testament, how this relates to the language of the Septuagint (for example, Rundgren 1965; Bentein 2011; Bentein 2013).[3] In this chapter, I blend these approaches to assess how multi-word formulations of the imperfect differ from their synthetic equivalents in meaningful ways. I argue that the difference triangulates an aspectual and indeed literary/narratological approach that is present in Luke's gospel, but only emerging in Mark's, offering three different ways to set up the background action in a scene.

Luke states his historiographical approach directly and also signals it with his specific phrasing. He opens the gospel with an unusual particle, ἐπειδήπερ (*epeidēper*) "when".[4] As Thrall (1962) documents, the New Testament writers make very little use of the richness available through Greek particles. Compounding particles is particularly weak, with even Luke using a statistically very low number of combinations (Thrall 1962: 8–9). Using such an unusual word, therefore, to open the text, sets our scene for a work that could be expected to be linguistically deliberate and rich. In verses 1-4, a preface dedicated to Theophilus, he establishes himself as wanting to write the story clearly, that is, to being in some sense a historian. Even phrases such as ἔδοξεν κἀμοὶ (*edoxen kamoi*) "it seemed a good idea to me" are reminiscent of Herodotus' ὡς ἔμοιγε δοκεῖ (*hōs emoige dokei*) "so it seems to me", casting Luke as an engaged and deliberate narrator, unlike in the other gospels.[5] Where Herodotus, however, uses this phrase to excuse himself from criticism over implausible content, Luke will do the opposite: take the thoroughly implausible story of Jesus' life, and commit to its veracity. With this deliberate positioning, it is appropriate to probe other ways in which Luke uses language to structure expression. Moreover, Herodotus (and other historians), themselves make specific, notable use of periphrastic constructions, highlighting the specific linguistic feature under discussion in this chapter (for example, Gonda 1959).[6]

3 The text of the New Testament dates to first and second centuries CE and that of the Septuagint, that is the translation of the Hebrew Old Testament into Greek, dates to about the third century BCE.

4 This is a New Testament *hapax legomenon*. The *Thesaurus Linguae Graecae* (https://stephanus.tlg.uci.edu/) gives 530 examples across the full Greek corpus.

5 The aorist version, as in Luke, is found seven times in Herodotus (*Histories*, 1.79; 1.172; 3.19; 3.64; 7.146; 8.64; 8.71; 9.86). See also Thrall (1962: 8) on Luke's positioning as a historiographer.

6 Probert (2018: 1) gives Herodotus, *Histories* 9.27.5 as her first example of a periphrastic construction.

2. Methodology

Classifying what is, or indeed is not, a periphrastic construction, is itself not uncontested.[7] Nardi (2025: 34), as a recent summariser, describes εἶναι *einai* "to be" + participle as the most studied and contested periphrastic construction. Nardi and Romagno (2025: 45) discuss "prototype" categories for these multi-word expressions, removing the focus on ascertaining precisely what kind of periphrastic construction may be present. I take such a broad view in this chapter. For the purpose of this chapter, I work within some specific parameters. My primary focus is on exploring the potential narratological impact and exegetical usefulness of a specific kind of multi-word expression in the imperfect. The impact of this is that precise classification of periphrasis is not essential to the argument. I use the term "multi-word expression" as the umbrella category of which periphrastic constructions are an example. When engaging with secondary literature, I echo other author's uses of the term periphrastic. This is, however, an argument about multi-word expressions broadly conceived, and their impact on understanding, not on the narrower category of periphrastic constructions and their definition. I explore these categorisation differences further below.

I include almost all imperfect verbs in the Greek New Testament. I use the term periphrastic with regards to the basic data as a shorthand to refer to any example of εἰμί (*eimi*, "I am") used with a participle in the same clause, that is, to any multi-word formulation with εἰμί (*eimi*) as the auxiliary verb. Bentein et al. (2012) discuss how existential uses of εἰμί (*eimi*) may be combined with participles, leading to a lack of clarity over whether an εἰμί (*eimi*) + participle combination is in fact periphrastic; their example is Luke 1:10. As they demonstrate, it is possible to read some examples in multiple ways (they categorise Luke 1:10 as both existential-progressive and periphrastic). In a separate study, Bentein (2013: 174) notes that 3/4 of his durative progressive examples from Mark and Luke are ambiguous; precise classification remains elusive. The scope of this chapter is not all periphrastic constructions in Mark and Luke, but all multi-word expressions using an imperfect auxiliary verb.

For the purpose of this chapter, I restrict discussion to εἰμί (*eimi*) as an auxiliary verb. Given my focus on the imperfect rather than periphrasis per se, I exclude examples with other tenses of εἰμί (*eimi*) from my main analysis but do comment briefly on these.[8] I include all participle tenses used with the imperfect of εἰμί (*eimi*); these are predominantly present or aorist, with occasional perfect participles.[9] Where more than one participle depends on the same auxiliary verb, in either Mark or Luke, I count it as "1.5" for multi-word-expression purposes.

7 Bentein (2011) explores some of the problems.

8 Mark 13:13, 13:25, 15:22, 15:34. Luke 1:20, 6:43, 7:8, 7:25, 12:2, 12:52, 14:8, 20:6, 21:17, 21:24, 23:15, and 24:38.

9 See below for further exploration of the range available.

I exclude οἶδα *oida* "to know", κάθημαι *kathēmai* "to sit", κείμαι *keimai* "to lie", and δύναμαι *dunamai* "to be able to" for the purpose of this chapter. I do this because they have limited past-tense forms, and so a clear differentiation between the imperfect and other tenses is less easy to formulate. The impersonal verb δεῖ (*dei*) "it is necessary", and its imperfect form ἔδει (*edei*), are also awkward in this context, given the lack of a full verbal morphology. I include examples in my basic statistics, but note that only Luke uses the imperfect form ἐδει (*edei*), adding weight to the sense that his use of the imperfect may be idiosyncratic. I also exclude the sole use of ἀπέκτεινεν (*apekteinen* "he killed") in the two gospels (Luke 13:4). As a liquid verb, the imperfect and aorist forms are indistinguishable. Given that killing is a fundamentally specific action rather than a durative one, I judge this to be aorist.

At Luke 17:27 and 17:28, (1), there are two "runs" of imperfect verbs:

(1) Luke 17:27-28

ἤσθιον, ἔπινον, ἐγάμουν, ἐγαμίζοντο, ... ἤσθιον, ἔπινον, ἠγόραζον, ἐπώλουν, ἐφύτευον, ᾠκοδόμουν·

ēsthion, *epinon*, *egamoun*, *egamizonto*... *ēsthion*, *epinon*, *ēgorazon*, *epōloun*, *efuteuon*, *ōikodomoun*.

"They were eating, drinking, marrying, getting married...they were eating, drinking, doing business, buying and selling, procreating, building houses."

Given that each is a collection of synthetic imperfects to set up an overarching scenario, I have treated each as a single instance. Where one auxiliary verb governs two participles, I have classified this as 1.5 uses of the imperfect. It is important to capture this variation, not least for instances where two different participle tenses are used within the same clause.

Having collated the examples of the imperfect in the New Testament, I then isolate the multi-word forms. Where both gospels share a story, I compare the expression in each. This intertextual approach is balanced by an intratextual one, wherein I consider how the difference between the synthetic and multi-word imperfect forms shape an individual gospel's narrative. I consider also the relationship between lexical meaning and morphological form, suggesting that some differences between synthetic and multi-word forms of the imperfect can be explained as lexically driven and aspectual.

3. Data

The New Testament is an artificial canon, with "New Testament Greek" a description of the idiolectical variation of that canon's authors rather than of

a particular Koine dialect (see Ryan (2024) for a more extensive discussion).[10] While it is therefore unhelpful to derive generalised linguistic principles from the New Testament, Luke's corpus (Luke-Acts), accounting for 25% of the text, is sufficiently large that we can establish some patterns in his writing.

The data was collected manually, given some of the morphological complexities involved in using most digital tools.[11] The text used for this article is the United Bible Society 5th edition, which is textually the same as the Nestle Aland 28th edition.[12]

	Mark	Luke
Total number of imperfects	263	328
Of these ἔλεγον *elegon* /ἔλεγεν *elegen* "I/they said; he said"	46	20
Total without ἔλεγον *elegon* /ἔλεγεν *elegen* "I/they said; he said"	217	308
Periphrastic imperfects	25	52
Periphrastic as % of total imperfects	9.51%	15.85%
Periphrastic as % of imperfects without ἔλεγον *elegon*/ἔλεγεν *elegen* "I/they said; he said"	11.52%	16.88%
Number of participles used in periphrastic imperfects	27	54
Of these in the present tense	20	39
Of these in the perfect tense	7	13
Of these in the aorist tense	0	2
Percentage of the present tense	74.07%	72.22%
Percentage of the perfect tense	25.93%	24.07%

10 Horrocks (2010) makes it one of four Koine branches, which remains vague, but still expects to see more consistency across the books and authors than may be appropriate.

11 A pilot analysis using Gemini Pro 3 was conducted. I collated imperfect verbs using Gemini Pro. I prompted Gemini to list all the imperfect verbs in Mark and Luke, based on the text by the United Bible Society 5th ed. / Nestle-Aland 28th edition. Initially, Gemini significantly undercounted these, identifying 214 in Mark, and 221 in Luke. Additionally, this included many examples that were not, in fact, imperfect verbs, but often, aorists, for example ἤρξατο *ērxato* "he began" (repeatedly). A later re-run of the task counted 507 in Mark, and 770 in Luke. This is a huge overcount based on my manual count of 263 and 328 respectively, even acknowledging I had removed a few specific verbs from the analysis. While Gemini could explain its methodology, it does not yet seem to be able to apply it sufficiently rigorously to the text to result in usable data. It was, however, useful for cross-checking purposes.

12 Word studies have been carried out using the *Thesaurus Linguae Graecae*, which draws on the text by United Bible Society 2nd edition (Aland 1968). There are inevitable discrepancies, but not to an extent that disrupts the argument.

	Mark	Luke
Percentage of the aorist tense	0.00%	3.70%
Other periphrastics	4	12

Table 8.1 Summary statistics of imperfect verbs and "periphrastic" constructions in the gospels of Mark and Luke.

Mark uses the imperfect form of λέγω (*legō*) "to speak" 46 times, accounting for 18.5% of his imperfects. Luke, in contrast, uses the imperfect of λέγω (*legō*) only 20 times, accounting for 6.1% of his uses of the imperfect. Given that this is a specific stylistic quirk in Mark, I therefore also use statistics excluding the use of the imperfect of λέγω (*legō*) from the analysis. I return to the difference towards my conclusion. Mark uses a multi-word version of the imperfect in 24.5 instances, accounting for 9.8% of his imperfect use including λέγω (*legō*), or 12.1% without. Luke, by contrast, uses the multi-word imperfect 14.1% of the time, or 15.0% without λέγω (*legō*).

While the raw numbers are similar, Luke's use is higher, sustained over a longer gospel, and reflects a different usage pattern. This chapter, therefore, builds an argument based on context, distribution, and function more than statistical difference.

4. Uses of the imperfect in Greek

To help understand what the multi-word imperfect phrases are doing differently to their synthetic counterparts, I first summarise some of the (relatively) agreed uses of the imperfect in Greek, especially in the New Testament Koine. Discussions over tense and aspect in Greek continue, with consensus yet to be agreed, but functional roles for the imperfect, however classified, can be broadly accepted.

Siebenthal (2019: 321–323) categorises the key difference between the imperfect and aorist (indicative) as aspectual, citing it as durative, and linear. Wallace (1996) describes the imperfect in the New Testament as giving an internal aspect, akin to a motion-picture experience from within the action, that is not so concerned with its beginning or end. This reflects Smyth's categorisation of it as representing an action as still going on, or a state still existing in the past (Smyth and Messing 2013: §1889). This progressive/descriptive sense, which also gives a vividness or simultaneity with other actions is what people often consider to be the imperfect's main role.[13]

13 See also Du Plooy (1991: 163) on uses of the imperfect. He gives six: progressive/durative/linear, conative, iterative, aoristic, moving panorama, and potential, with specific reference to the New Testament, but without examples from Mark or Luke. Du Plooy (1991: 166–168) explores

This is the usual "was doing" kind, or as Smyth writes: "The imperfect is often used to describe an action or state that is in progress in past time from the viewpoint… of the speaker.… It speaks either of *vividness* or *simultaneity* with another action" (Smyth and Messing 2013: §1898).[14] While the sense of being *in medias res* is clear, the inchoative force of the imperfect should not be dismissed, and remains relevant to the New Testament.[15] As Robertson (1914: 885) notes, the inchoative (or inceptive) imperfect is particularly useful for describing the 'beginning of an action' rather than its mere occurrence, a feature that aligns with the introductory nature of Jesus's teaching in the early chapters of Mark and Luke.

There are also times when the imperfect will be instantaneous, that is, similar to the aorist. These mainly involve ἔλεγεν (*elegen*) as "he said", particularly in Mark (46 uses), but also for example, Matthew 9:24, Luke 23:42, John 5:19. The imperfect can, with ἔλεγεν (*elegen*), be a stylistic tool to keep a narrative flowing, while still marking points of action and change in the narrative akin to the aorist (see Smyth and Messing 2013: §1908; Siebenthal 2019: 323).

The imperfect may also be used iteratively (where Wallace's standard examples do not include Mark or Luke), or conatively, to suggest attempted action which may or may not have been completed (for which Wallace suggests Matt 3:14, Mark 15:23, Acts 26:11, Luke 1:59, Rom 9:3.).[16] It can be used in a customary or habitual way, for which Wallace offers Matt 26:55, Mark 4:33, Luke 2:41, Acts 3:2, Rom 6:17, Gal 1:14. It may also be used pluperfectively, to indicate antecedent action relating to a main narrative expressed in the past. Wallace (1996) suggests that this is more common with ἔλεγεν (*elegen*), offering Mark 5:8 and 6:18, and Luke 8:29 as examples (see also Smyth and Messing 2013: §1906).[17]

In summary, the imperfect is a versatile verb form which was used in a broad range of ways to describe actions within both tense and aspect frames. Reducing this to a simple guide has eluded grammarians. We can complicate the picture further by considering the multi-word, or periphrastic version.

aspect in Luke-Acts more specifically, noting Luke-Acts 1:22 as periphrastic, emphasising the periphrastic as a marker of continuation.

14 Wallace (1996) also cites Fanning (1987: 241).

15 The terms inchoative, ingressive, and inceptive all describe the use of the imperfect to imply an action's beginning.

16 See Wallace (1996: 550–551), Siebenthal (2019: 322), and Smyth and Messing (2013: §1895) on the conative imperfect. For a fuller discussion about conatives, see Malone (2020).

17 In addition, Smyth and Messing (2013: §1890) classify an imperfect of continuance, and Smyth and Messing (2013: §1901), the imperfect in place of a present.

5. The multi-word imperfect

The synthetic imperfect, for example, ἐδίδασκεν (*edidasken*) "he was teaching", is also accompanied, throughout the New Testament, by a multi-word version, for example, ἦν διδάσκων (*ēn didaskōn*) "he was teaching". This multi-word expression formulation of the imperfect, pairing an auxiliary verb (εἰμί *eimi*) with a participle, is relatively frequent in the gospels. In his dissertation on the grammaticalisation of time and use of the imperfect in the New Testament, Ortner (2019: 62) makes only one passing reference to periphrasis (as a specific form of multi-word expression. Johnson (2010) gives an extensive account, but makes only one reference to Luke. He suggests that the periphrastic imperfect has a discourse function, but is limited to narrative sections of the gospel, and that its use is over-reported. He argues that it has been under-investigated because of its assumed grounding in Semitic grammar, but that it represents, instead, a more marked form of Koine.[18] His analysis is limited, and while I return to it, I acknowledge Bentein's (2013: 170) criticism of the narrowness, and in this chapter use a broader definition of multi-word expression.

Daniel (2021: 729) explores how frequent ἦν (*ēn*) with a present participle is in the Septuagint (see also Bentein 2013). Bentein (2016: 248) observes that as many as 82% (146/178) of the examples of what he identifies as periphrastic constructions in the New Testament are, consequently, structurally influenced by their Hebrew parent text.

Perhaps the best-known example of the periphrastic/multi-word imperfect in the gospels is ἦν διδάσκων (*ēn didaskōn*) "he was teaching". This occurs seven times in the New Testament, five of which are in Luke.[19] It becomes a stylistic quirk of Luke's, but in this chapter, I suggest that it is more than idiolectical. It is to some degree intentional, contrasting with both the aorist and the synthetic imperfect as a way to introduce scenes. I turn now, therefore, to compare its use with the synthetic equivalent, ἐδίδασκεν (*edidasken*).

ἐδίδασκεν (*edidasken*) is used 14 times in the New Testament. Six are in Mark (1:21, 2:13, 4:2, 9:31, 10:1, 11:17). Three are in Luke-Acts (Luke 4:15, 5:3, and Acts 2:14). This weighting reverses when we consider the multi-word version ἦν διδάσκων (*ēn didaskōn*), used once each in Matthew and Mark (Matthew 7:29, Mark 1:22), but five times in Luke. The sole use in Mark comes just one verse after its synthetic equivalent, (2).

(2) Mark 1:21-22

21 Καὶ εἰσπορεύονται εἰς Καφαρναούμ· καὶ εὐθὺς τοῖς σάββασιν εἰσελθὼν εἰς τὴν συναγωγὴν ἐδίδασκεν. 22 καὶ ἐξεπλήσσοντο ἐπὶ τῇ διδαχῇ αὐτοῦ· ἦν γὰρ διδάσκων

18 See also Nardi (2025) on the over-focus on Semitic influence.
19 ἦν διδάσκων *ēn didaskōn* appears 5 times in Luke, at 4:31, 5:17, 13:10, 19:47, 21:37.

αὐτοὺς ὡς ἐξουσίαν ἔχων καὶ οὐχ ὡς οἱ γραμματεῖς.

21 Kai eisporeuontai eis Kafernaoum; kai euthus tois sabbasin eiselthōn eis tēn sunagōgēn edidasken. 22 kai exeplēssonto epi tēi didachēi autou; ēn gar didaskōn autous hōs exousian echōn kai ouch hōs hoi grammateis.

"They were journeying into Capernaum. Immediately on the Sabbath, after he'd entered the synagogue, he began to teach. Everyone was astounded at his teaching. For he was teaching them as though with authority and not like the scribes."

The synthetic imperfect makes sense as an inchoative here, completing the sequence of actions that precipitates the ongoing astonishment. The multi-word version, with its weak explanatory particle γάρ (*gar*) "for", shifts the focus onto the novelty of the teaching, rather than just its occurrence. Luke makes a similar statement at 4:32, (3), following a multi-word expression, and preceding a miracle:

(3) Luke 4:32

καὶ ἐξεπλήσσοντο ἐπὶ τῇ διδαχῇ αὐτοῦ, ὅτι ἐν ἐξουσίᾳ ἦν ὁ λόγος αὐτοῦ.

kai exeplēssonto epi tēi didachēi autou, hoti en exousia ēn ho logos autou.

"They were struck by his teaching, because his speech was using authority."

Where Mark has used the multi-word phrase in the context of astonishment and authority, Luke foregrounds the teaching even further by using the noun itself, for the only time in his gospel. I suggest that this is the pattern we see further developed in Luke. Luke uses the synthetic third-person singular form just twice (4:15, 5:3). These are both early on in the gospel, before the sense of never-ending teaching to which I suggest the multi-word imperfect is indicating has started. Levinsohn (2016a: 170) argues for a tension between a stative and a background progressive use. Van Emde Boas et al. (2019: §33.15–17) identify the imperfect as the primary vehicle for "progressive" aspect in past time, often providing the "frame" for other events. By choosing the multi-word form, Luke emphasises this framing function over the simple process of the verb. In comparing the synthetic and multi-word examples both between and within the gospels, I suggest that we see a development between these options.

Even in the early chapters, Luke's use contrasts directly with Mark's, including where we still read a synthetic equivalent, (4).

(4) Luke 4:31

Καὶ κατῆλθεν εἰς Καφαρναοὺμ πόλιν τῆς Γαλιλαίας. καὶ ἦν διδάσκων αὐτοὺς ἐν τοῖς σάββασιν·

Kai katēlthen eis Kapharnaoum polin tēs Galalaias. kai ēn didaskōn autous en tois sabbasin.

"He went into Capernaum, a city in Galilee. He was teaching/used to teach there

on the Sabbath."

This contrasts with Luke 4:15, (5).

(5) Luke 4:15

καὶ αὐτὸς ἐδίδασκεν ἐν ταῖς συναγωγαῖς αὐτῶν, δοξαζόμενος ὑπὸ πάντων.

kai autos edidasken en tais sunagōgais autōn, doxazomenos hupo pantōn.

"He was teaching in their synagogues, glorified by everyone."

Here, the synthetic form potentially demonstrates a general teaching episode, not one where anything marked or unusual needs to be noted. The teaching itself is not the key point, whereas later in the gospel, teaching becomes more marked given its content, authority, and timing (for example, on a Sabbath), (6).

(6) Luke 5:3

καθίσας δὲ ἐκ τοῦ πλοίου ἐδίδασκεν τοὺς ὄχλους.

kathisas de ek tou ploiou edidasken tous ochlous.

"After sitting down on the boat, he began to teach the crowds."

This prefaces his call to others to join him as disciples, and so truly is marking the start of his teaching mission, and so again, an inchoative force seems appropriate.[20] This contrasts with Luke 5:17, (7).

(7) Luke 5:17

Καὶ ἐγένετο ἐν μιᾷ τῶν ἡμερῶν καὶ αὐτὸς ἦν διδάσκων

Kai egeneto en miai tōn hēmerōn kai autos ēn didaskōn

"It happened on one of the days that he was teaching…"

The multi-word equivalent here leads into a collection of similar constructions discussed further below. At this point, its contrast with Luke 5:3 is important; it heralds a miracle, and a debate about the relationship between teaching through words versus actions. The two early uses of ἐδίδασκεν (*edidasken*) "he

20 The inchoative sense of the imperfect could also be communicated using ἄρχομαι (*archomai*) with an infinitive. The opening of Luke-Acts provides a good contrast here: Luke-Acts 1:1 Τὸν μὲν πρῶτον λόγον ἐποιησάμην περὶ πάντων, ὦ Θεόφιλε, ὧν ἤρξατο ὁ Ἰησοῦς ποιεῖν τε καὶ διδάσκειν, *ton men proton logon epoiēsamēn peri pantōn, ō Theofile, hōn ērxato ho Iēsous poiein te kai didaskein* "I have done the first tale about everything, Theophilus, that Jesus began to do and teach." Here, the aorist of ἄρχομαι (*archomai*) is used with the infinitive of διδάσκω *didaskō*). This demarcates a break between the two texts; while they continue a history, it is no longer the ongoing dialogue expressed by either ἐδίδασκεν (*edidasken*) or ἦν διδάσκων (*ēn didaskōn*).

was teaching", I suggest, play a specific role in beginning Jesus' teaching mission, but once it is established, and where it is novel or controversial, the multi-word expression becomes more appropriate.

There is a sense in Luke's gospels, of an undercurrent of constant dialogue with the Pharisees, of which only small elements "break through" into the text to be witnessed in a narrative. This accords with a further sense in Luke that central chapters represent an ongoing dialogue/debate between Jesus and the Pharisees, only some of which is captured in the text.[21] The use of ἀποκρίνομαι (*apokrinomai*) "to answer, to reply" is important here.[22] Both the aorist participle, ἀποκριθείς (*apokritheis*), and indicative, ἀπεκρίθη (*apekrithē*) occur frequently. ἀποκριθείς (*apokritheis*) is particularly Lukan, with 34/94 uses of the participle in the New Testament being from Luke-Acts; the relational quality of his writing is made clear. Sometimes this is in genuine reply to someone, for example Luke 10:27, where Jesus replies to the lawyer. At other times, it is to a situation, for example at Luke 11:14 where the high priest responds to Jesus healing on the Sabbath.

There remain times, however, where the provocation for response is less clear; at Luke 14:3 there is an assumption that Jesus will heal the man, and cause consternation as a result. His reply is almost pre-emptive of what we have seen in places like Luke 11:14. The use of ἀποκρίνομαι (*apokrinomai*) "to answer, to reply" provides a way to show that the gospel is not the whole teaching, and the ongoing murmuring against Jesus is a constant conversation. This point is reinforced by Luke's use of διαγογγύζω (*diagogguzō*) "to murmur", used only twice in the New Testament, both by Luke, in the imperfect, of the Pharisees and those suspicious of Jesus (Luke 15:2, 19:7). Luke is the only author to use ἐκμυκτερίζω (*ekmukterizw*) "to mock", both in the imperfect (Luke 16:14, 23:35). He is the only author to use the imperfect of πυνθάνομαι (*punthanomai*) "to enquire" (and 9/12 total uses are Luke-Acts). Similarly, although there are 56 uses of ἐπερωτάω (*eperōtaō*) "to ask" in the New Testament, only Mark and Luke use it in the imperfect, six times in Mark, but only four in Luke. The verb ζητέω (*zēteō*) "to seek" occurs 120 times in the New Testament, 35 times in Luke-Acts, which is roughly proportionate, but in the imperfect, it is 13/28 uses. I suggest that this focus on enquiry and response as ongoing, iterative, imperfect actions is indicative of a pattern which is hinted at in Mark, particularly with his frequent use of ἔλεγεν (*elegen*), but developed in Luke.

21 See Kilgallen (1997) on the conversational continuity between Luke 15 and 16, for example.

22 I cite it using its New Testament middle/passive deponent form, rather than its more classical active form ἀποκρίνω (*apokrinō*).

5.1 καὶ ἐγένετο *kai egeneto*: Aorist scene-setting as a contrast

Example (7) includes not only an imperfect setting, but an aorist opening: καὶ ἐγένετο (*kai egeneto*) "and it happened". The grammar of scene-setting is relevant to both pericopes and chapters. Half of the chapters in Luke open with an imperfect (as opposed to only a third in Mark, and only two chapters in Matthew).[23] Luke, however, uses ἐγένετο (*egeneto*)/ καὶ ἐγένετο (*kai egeneto*) 68 times. Although it is often seen as a Markan phrase, it occurs only 17 times in Mark.

Luke's use is copious, and not purely formulaic. Reilig (1965: 159–160) suggests that Luke is using it not as a marker of a majestic or hieretic style, reflecting the Hebrew phrasing of the Old Testament, but as a development of the Greek of the Septuagint. While the Septuagint is itself a reflection of the Hebrew, it is also its own form of Greek, with Atticising tendencies, and other cultural influences. Luke is primarily a Hellenist, rather than a Semiticist. His use of καὶ ἐγένετο (*kai egeneto*) is distinctive within Greek, with its own "scene-setting" impact. Five of Luke's uses include specific time references (Luke 1:5, 1:59, 2:1, 17:26, 17:28), while 22 are followed by an articular infinitive to give a durative background. This includes both Luke 9:18 and 9:28, (8).

(8a) Luke 9:18

Καὶ ἐγένετο ἐν τῷ εἶναι αὐτὸν προσευχόμενον

Kai egeneto en tōi einai auton proseuchomenon

"and it happened that while he was praying…"

(8b) Luke 9:28

καὶ ἐγένετο ἐν τῷ προσεύχεσθαι αὐτὸν

kai egeneto en tōi proseuchesthai auton

"and it happened that while he was praying…"

Both of these could be translated as "and it happened that while he was praying…", but Luke 9:18 uses a multi-word (periphrastic) construction with the infinitive, while Luke 9:28 uses a simple infinitive.[24] Nardi (2025: 45) uses this difference to suggest that there is no functional difference between the synthetic and periphrastic forms. I suggest that the two uses are, instead, not intended to be comparable, and there is a distinction; as Nardi goes on to discuss, "difference" may come in different categories, including pragmatic,

23 See Levinsohn (2016b: 314) on the use of the imperfect periphrastic to begin pericopes.
24 See Levinsohn (2016b: 312) commenting on Turner's description of these two examples.

semantic, or aspectual. While not imperfect multi-word expressions, they do show clearly the different ways in which Luke sets up scenes using a range of temporal and aspectual markers. This combination of aorist frame and temporal location through an infinitive provides a contrasting way to establish a scene that is distinctive to Luke; only one of Mark's uses of καὶ ἐγένετο (*kai egeneto*) is similar (Mark 4:4).

Developing this further, Bentein's (2013: 184) fourth class of focalised progressive periphrastic notes only two examples, both from Luke (5:17–18 and 14:1). Both have a pattern of καὶ ἐγένετο (*kai egeneto*) followed by a temporal use of a gerund, and then a progressive periphrastic, (9).

(9) Luke 5:17–18

Καὶ ἐγένετο ἐν μιᾷ τῶν ἡμερῶν καὶ αὐτὸς ἦν διδάσκων, καὶ ἦσαν καθήμενοι Φαρισαῖοι καὶ νομοδιδάσκαλοι οἳ ἦσαν ἐληλυθότες ἐκ πάσης κώμης τῆς Γαλιλαίας καὶ Ἰουδαίας καὶ Ἰερουσαλήμ· καὶ ἰδοὺ ἄνδρες φέροντες ἐπὶ κλίνης ἄνθρωπον ὃς ἦν παραλελυμένος.

Kai egeneto en miai tōn hēmerōn kai autos ēn didaskōn, kai ēsan kathēmenoi Farisaioi kai nomodidaskaloi hoi ēsan elēluqotes ek pasēs kōmēs tēs Galilaias kai Ioudaias kai Ierousalēm; kai idou andres ferontes epi klinēs anthrōpon hos ēn paralelumenos.

"It happened on one of the days he was teaching, and the Pharisees were sitting down, along with the teachers of the law who had come from every village of Galilee, and Judaea, and Jerusalem. See, men carrying a man on a bed who was paralysed."

Luke's use of καὶ ἐγένετο (*kai egeneto*) is already idiosyncratic, and with this further multi-word imperfect focus, we see a specific approach to scene-setting. The aorist situates us in a new moment, while the imperfect expresses the dynamic nature of inhabiting that moment. In this case, we find four periphrastic constructions within two verses, each with a different temporal focus. Bentein is interested in only the last of these, but the pattern with καὶ ἐγένετο (*kai egeneto*) is important in general, (10).

(10) Luke 14:1

Καὶ ἐγένετο ἐν τῷ ἐλθεῖν αὐτὸν εἰς οἶκόν τινος τῶν ἀρχόντων [τῶν] Φαρισαίων σαββάτῳ φαγεῖν ἄρτον καὶ αὐτοὶ ἦσαν παρατηρούμενοι αὐτόν.

Kai egeneto en tōi elthein auton eis oikon tinos tōn archontōn [tōn] Farisaiōn sabbatōi fagein arton kai autoi ēsan paratēroumenoi auton.

"It happened that while he was going into the house of one of the leaders of the Pharisees on the Sabbath to eat bread, they were watching him closely."

These two examples illustrate several key points about what makes the multi-

word imperfect different in Luke's gospel.

5.2 Multi-word imperfects in Luke

This sense that there was teaching taking place in a generic but important way helps explain other observations made about the clause, and integrate some apparent differences of opinion. Gonda (1959: 97) opens by countering Rosén's (1957) argument that periphrastic clauses occur when the idea of a verb is not the predicate of a sentence, but the subject, and the logical predicate is not the agent, but a complement. As Gonda explains, verbal sentences have tended to be used for actions, while nominal/semi-nominal ones are used for description. With this development of periphrastic clauses, we see a move towards more nominal sentences being used to set up action. Gonda (1959: 99–100) suggests that whereas the predicate of a verbal sentence focusses the hearer's thoughts on what happens, the predicate of a nominal sentence directs their attention to what is (i.e. more existential). With ἦν διδάσκων (*ēn didaskōn*), the important issue, on this model, would be that teaching were taking place, moreover in the synagogue, and on a Sabbath, rather than necessarily it being Jesus himself doing the teaching. Jesus' agency can be assumed; this is, after all, his story. The novelty of what he is doing (where, when, how), cannot. By using the multi-word imperfect, therefore, on Rosén's model, we shift the focus onto the novelty of the teaching, and see this borne out in the unusual linguistic use of other terms around it. This does not require it to be a "true" periphrastic construction in any of the defined senses; the multi-word expression still stands in contrast to the synthetic equivalent.

σαββάτῳ (*sabbatōi* "on the Sabbath"), is another irregularity demarcating difference. It appears in both the singular and plural. When singular, as quoted above, it is often without a definite article, giving a generic sense of "on a Sabbath, any old Sabbath". When plural, it would usually be construed of as "week" (i.e. seven days), but clearly refers to the Sabbath, and so seems to carry a sense, again, of "pick a Sabbath, any Sabbath", because it is not the date that matters so much as the context of what teaching in a synagogue on a Sabbath means, as opposed to anywhere else, at any other time (cf. Levinsohn 2016b: 313). Luke has both maintained the novelty of Jesus' Sabbath teaching, and the ongoing undercurrent of teaching and enquiry to which I suggest the imperfect points.

Rosén's approach, however, is not the only way to understand the multi-word (and sometimes periphrastic) imperfect, particularly with reference to ἦν διδάσκων (*ēn didaskōn*). Gonda (1959: 107) sees Rosén as overreaching himself, and the emphasis of one participant in the construction over another as too subtle. I suggest not that Rosén is overreaching/overly subtle, but that his

understanding of difference between synthetic and periphrastic construction is in part grounded in both aspectual and narratological differences.[25] Bringing in other perspectives, and both literary and lexical arguments gives Rosén's analysis a role integrated with the work of others, notably Björck, Bentein, and Johnson.

Björck (1940) takes a very different approach, with specific reference to the New Testament. He identifies improper periphrasis, where the participle is not a predicate, and εἶναι (*einai*) "to be" is not a copula, adjectival periphrasis, where the participle has the character of an adjective, and "real" periphrasis.[26] Björck (1940) cites Luke 9:47 and ἦν διδάσκων (*ēn didaskōn*) as characterising "real" periphrasis, where the process is being conceived as a temporal frame encompassing something else which may be expressed but may also be understood from the context (see also Gonda 1959: 104). It becomes about the relative duration of processes. The imperfect in general establishes ongoing background contextual action across which specific moments are overlaid. Levinsohn's (2016b: 312) variation on this, wherein what he calls the "copular imperfect" presents iterative events, portraying an actor as performing an action from time to time during a period rather than continuously may also point towards the reading I am giving; the multi-word imperfect helps to set up a situation wherein what is represented in the text is a "breakthrough" from a background situation, with iterative but distinct examples of an action, such as teaching. Taking Rosén's view, the multi-word version of it may help shift attention from agent to action. Blending these together, we see how the multi-word imperfect focusses our attention not only on action, but on the durative or progressive nature of that action. Every ἦν διδάσωκν (*ēn didaskōn*) therefore also becomes a cumulative contribution to the overall "didache" (teaching).

Beyond this, there are other ways in which Luke's use of what Bentein identifies as progressive periphrastic constructions are unusual. These include:

25 Nardi and Romagno (2025: 46) cite Luke 4:20 as their sixth category of periphrastic construction, linked to aspectual variation. Nardi and Romagno (2025: 50) discuss aspectual differences in periphrastic constructions, but across too broad a range of texts to be directly relevant to a New Testament study.

26 See Gonda (1959: 105) for a summary. This adjectival use is clearly true of Luke, see Bentein, Janse, and Soltic (2012), for example, and for a Greek example, see Luke 15:24 (echoed in Luke 15:32): ὅτι οὗτος ὁ υἱός μου νεκρὸς ἦν καὶ ἀνέζησεν, ἦν ἀπολωλὼς καὶ εὑρέθη *hoti houtos ho huios mou nekros ēn kai anezhsen, ēn apolōlōs kai heurethē* "Because my son here was dead and has come alive, was lost and is found", where "dead" and "lost" occupy identical grammatical positions in the description, but one is an adjective, the other a participle. Chapter 15, however, is also heavily focussed on ways of using ἀπόλλυμι (*apollumi*) "to destroy", with 8/29 of Luke's uses of the verb (90 in the New Testament) in just that one chapter (Luke 15:4, 15:4, 15:6, 15:8, 15:9, 15:17, 15:24, 15:32). These verses are unexpected for other reasons and would need further analysis beyond the scope of this chapter.

1. ἰδού (*idou*) as the focalisation point for a periphrastic progressive (Luke 13:10—ἦν διδάσωκν (*ēn didaskōn*), Acts 1:10, 10:30, 12:6) (Bentein 2013: 184–185).[27]
2. The periphrastic progressive in a subordinate clause preceded by καὶ ἐγένετο (*kai egeneto*) (Luke 9:18, 11:1) (Bentein 2013: 186).[28]
3. The periphrastic progressive occurs before its focalisation point in a subordinate clause (Luke 5:29) (Bentein 2013: 188).
4. The progressive construction in a subordinate clause following a foregrounded event serving as a focalisation point (Acts 2:2, 12:12) (Bentein 2013: 189).

Luke's use of the construction is more varied and nuanced than any other New Testament writer, including within the specific realm of the imperfect. Levinsohn (2016b: 314) summarises the imperfect periphrastic as describing ongoing states, presenting iterative events, and background scene-setting events at the beginning of pericopes. In what I have presented, I have suggested a way in which these functions blend in Luke, and become dependent on lexical context (verbs of enquiry/teaching etc.), offering a specific narratological function grounded in historiographical awareness.

5.3 Comparing Mark and Luke

There are places where Mark and Luke tell the same story, for example the healing of the Gerasene (Gadarene) Daimoniac (Matthew 8:28–34, Mark 5:1–20, Luke 8:26–39). Matthew omits the detail, but both Mark and Luke include a description of herd of pigs grazing on a mountain, (11).

(11a) Mark 5:11

῾Ην δὲ ἐκεῖ πρὸς τῷ ὄρει ἀγέλη χοίρων μεγάλη βοσκομένη

Ēn de ekei pros tōi orei agelē choirōn megalē boskomenē

"There was, there on the mountain, a big herd of pigs, grazing."

(11b) Luke 8:32

῾Ην δὲ ἐκεῖ ἀγέλη χοίρων ἱκανῶν βοσκομένη ἐν τῷ ὄρει·

Ēn de ekei agelē choirōn hikanōn boskomenē en tōi orei;

"There was, there on the mountain, a herd of suitable pigs, grazing."

27 Blending two of his categories, where he splits examples depending on clause order.
28 Note that neither of these examples uses the imperfect of εἰμί (*eimi*).

In both cases, the participle is separated from the auxiliary verb by a considerable distance (Bentein 2011: 14, refuting Porter 1992). The extent to which this matters is debated, but it may suggest that we should read the ἦν *ēn* "there was" more existentially, with the participle βοσκομένη (*boskomenē*) "grazing" adjectivally, rather than within a periphrastic context.[29] Two differences between Mark and Luke's versions may be important, however. In Mark's version, the participle follows closely after the noun ἀγέλη *agelē* "herd", the genitive describing it (χοίρων *choirōn* "of pigs") and an adjective μεγάλη *megalē* "big". It reads as part of that noun-phrase. In Luke, however, the removal of μεγάλη *megalē* "big" breaks that sense of continuity and leaves the participle more exposed. In addition, the phrase ἐν τῷ ὄρει *en tōi orei* "on the mountain" appears after the participle, rather than between the auxiliary verb and the participle, as in Mark. This removes some of the material that Porter (1992: Chapter 1, section 3.1) might argue reduced the potential for reading the phrase as periphrastic. The two points together bring the participle four words closer to the auxiliary verb, as well as reframing it. This makes it much easier to read the clause as a multi-word expression, as opposed to an existential use of the copula with an adjective. It is not a definitive point, but it potentially demonstrates ways in which Luke, in identical circumstances, uses multi-word expressions differently.

This difference also applies to his tense usage within the constructions more generally; Luke uses a broader range of tenses in his multi-word constructions than does Mark. Levinsohn (2016b: 311) determines the tense of a periphrastic construction from the tense of the participle, not the copula.[30] In this chapter, I have taken the opposite approach. My focus is the imperfect in general, but also the ways in which it can be targeted and nuanced through combination with different tenses of participles. In the imperfect, both Mark and Luke predominantly use the present participle (Mark 74%, Luke 71%), with a significant remaining number of aorist participles (Mark 26%, Luke 25%), but in addition, 4% of Luke's uses (as in Luke 5:17), are perfect participles.[31] In the multi-word constructions using other tenses of εἰμί (*eimi*), all four examples in Mark use present participles, whereas six of the twelve in Luke use perfect participles (Luke 12:2, 12:52, 14:8, 20:6, 23:15, 24:38). These constructions give Luke a way to place things both temporally and aspectually with more care than we see in Mark, and with a kind of perfectivity not seen elsewhere in the New Testament.

29 See Rundgren (1965: 467) on this phrase as related to the Semitic periphrastic construction.

30 See Nardi (2025: 38–40) on tenses in participles, and Nardi and Romagno (2025: 50) on tense distributions across time.

31 Levinsohn (2016b: 307, esp. n. 2) only allows either present or perfect participles. He cites Rijksbaron (2006) on the construction using an aorist participle, but makes these adjectival rather than periphrastic. He cites Porter (1992) and Campbell (2009) in agreement of his position. Given my broader study of "multi-word expression" beyond any strict interpretation of "periphrastic", I include the aorist examples.

Mark's more stative use of the multi-word imperfect is in keeping with his less literary style; periphrastic constructions may move from stative to progressive, to habitual.[32] Bentein has argued for Luke's progressive use. With this sense of an ongoing narrative into which phrases like ἦν διδάσκων (*ēn didsakōn*) enable us to dip, we see a move further along this spectrum to something that is potentially more habitual. This makes sense within the context of Luke's general use of both multi-word expressions, and the imperfect. There are eighteen uses of the imperfect of μέλλω (*mellō*) "to be about to" in the New Testament, for example, seven of which are in Luke-Acts (and none in Mark), a disproportionate number.[33] It shows us how carefully Luke uses imperfect verbs within multi-word expressions to demarcate and locate his actions precisely.

What we see in Luke, therefore, is a careful balance of καὶ ἐγένετο (*kai egeneto*), particularly with the articular infinitive), a synthetic imperfect, and a multi-word imperfect, to demarcate different kinds of scene. Luke may not necessarily use the imperfect tense more often than Mark, but he does use it in different ways, elements of which are seen in Mark, but which Luke develops much further. We find, in particular, a pattern of enquiry and discussion expressed using the imperfect, and movements in narrative demarcated using the imperfect, often as a way to reorient us within this ongoing discussion with the Pharisees. With Luke's use of the multi-word imperfect, I suggest we see a further development of this, offering three different ways to orient a reader in action: aorist, synthetic imperfect, multi-word imperfect.

6. Conclusion

In this chapter, I aimed to blend an understanding of how the imperfect functions for scene-setting purposes in the Greek New Testament, with an understanding of how a multi-word expression of the imperfect differs from a synthetic one. In so doing, across the gospels of Luke and Mark, I have demonstrated a narratological and exegetical difference between the formulations, and indeed, between the gospels. While Mark and Luke use the imperfect tense a broadly similar amount of the time, they do different things with it; in some places, Luke develops an idea already seen in Mark, while in others, his use is more contrasting. Mark's use of ἔλεγεν (*elegen*) is distinctive, and has long been recognised as such. What has been less appreciated, however, is the extent to which Luke specifically uses the imperfect to set scenes within a framework involving discussion. Luke,

32 Bertinetto, Ebert, and Groot (2000) discussed throughout Bentein (2013); their paper is less directly relevant to Greek, and not directly cited in this chapter. See also Nardi and Romagno (2025: 51) for data and discussion about the stative use of these phrases.

33 Ten are John/Revelation. Revelation is highly eschatological, and often uses periphrastic ways of expressing the future in place of a synthetic future tense.

moreover, uses the multi-word expression of the imperfect to set up a scene where action matters, or describe a state of being, whereas Mark uses it more for describing specific conditions rather than background action.

In addition, Luke makes greater use of the tenses of participles available, particularly pairing an imperfect auxiliary verb with a perfect participle, to describe the results of past actions that continue to be relevant. With this further observation regarding Luke's unusual use of the perfect participle in multi-word clauses, we return to Herodotus, and Luke the historiographer. Gonda (1959: 112) suggests that part of the reason the periphrastic construction developed in Greek was the increased downplaying of the pluperfect. She argues that in order to maintain a vivid narrative, periphrasis began to be needed. Additionally, the periphrastic construction allows an author to modify any idea expressed by the copula more clearly. Nardi (2025) extends this diachronic assessment by considering the relative frequencies of participle tenses through time, and the impact of the Old Testament on that distribution. Luke's use of the perfect participle reflects an older style, further underpinning our understanding of his language as more Hellenic, and indeed Atticising or archaising than other gospel writers.

Given what I have demonstrated about both the imperfect and the multi-word construction in Luke, what we may be seeing, therefore, is an author who archaises (Atticising, particularly with a view to historiography), integrates and synthesises (absorbing Latin and Semitic strategies into Greek), and innovates. This innovation takes several forms. Luke's use of the multi-word imperfect moves us into an ongoing dialogue with the didache, curating our journey through the lessons of the New Testament. In using multiple participle tenses, he situates the actions more carefully, and balances both stative and progressive concepts with more precision. Like a good cinematographer, Luke draws our attention to different elements of the story, through different people's eyes and perspectives. By framing Jesus's teaching through the multi-word imperfect, Luke suggests that the didache is not merely a series of historical incidents, but a continuous, backgrounded reality. Through this aspectual foregrounding, the multi-word construction functions as a tool of extradiegetic framing, stretching the narrative duration to encompass the reader in the ongoing dialogue. The specific events—the healings and the parables—are thus presented not as isolated snapshots but as "breakthroughs" into a persistent divine discourse.

Acknowledgements

Particular thanks go to Jonathan Middleton, who helped me extensively with this project.

Bibliography

Aland, Kurt. 1968. *The Greek New Testament*. London: United Bible Societies.

Bentein, Klaas. 2011. Towards the Identification of Verbal Periphrasis in Ancient Greek: A Prototype Analysis. *Acta Classica* 54. 1–25.

— 2013. The Syntax of the Periphrastic Progressive in the Septuagint and the New Testament. *Novum Testamentum* 55(2). 168–192. https://doi.org/10.1163/15685365-12341410

— 2016. *Verbal Periphrasis in Ancient Greek: Have- and Be- Constructions*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198747093.001.0001

Bentein, Klaas, Mark Janse, and Jorie Soltic. 2012. "And the Mass Was Praying Outside": A Note on Luke 1:10. *Neotestamentica* 46(1). 1–8.

Bertinetto, Pier Marco, Karen H. Ebert, and Casper Groot. 2000. The Progressive in Europe. In Östen Dahl (ed.), *Tense and Aspect in the Languages of Europe*, 517–558. Berlin: Mouton de Gruyter.

Björck, Gudmund. 1940. *ΗΝ ΔΙΔΑΣΚΩΝ: Die periphrastischen Konstruktionen im Griechischen*. Leipzig: Otto Harrassowitz.

Campbell, Constantine R. 2009. *Basics of Verbal Aspect in Biblical Greek*. Grand Rapids, MI: Zondervan.

Daniel, Andrew Glen. 2021. The Translator's Tell: Translation Technique, Verbal Syntax, and the Myth of Old Greek Daniel's Alternate Semitic Vorlage. *Journal of Biblical Literature* 140(4). 723–749. https://doi.org/10.15699/jbl.1404.2021.5

Du Plooy, Gerhard P. V. 1991. Aspect and Biblical Exegesis. *Neotestamentica* 25(1). 157–170.

Emde Boas, Evert van, Albert Rijksbaron, Luuk Huitink, and Mathieu de Bakker. 2019. *Cambridge Grammar of Classical Greek*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781139027052

Fanning, Buist M. 1987. A Study of Verbal Aspect in New Testament Greek. DPhil thesis (University of Oxford).

Gonda, Jan. 1959. A Remark on "Periphrastic" Constructions in Greek. *Mnemosyne* 12(2). 97–112.

Horrocks, Geoffrey C. 2010. *Greek: A History of the Language and its Speakers*. 2nd edn. Chichester: Wiley Blackwell. https://doi.org/10.1002/9781444318913

Johnson, Carl. 2010. A Discourse Analysis Of The Periphrastic Imperfect In The Greek New Testament Writings Of Luke. PhD Thesis (University of Texas at Arlington), https://mavmatrix.uta.edu/linguistics_tesol_dissertations/8

Kilgallen, John J. 1997. Luke 15 and 16: A Connection. *Biblica* 78(3). 369–376.

Levinsohn, Stephen H. 2016a. Verb Forms and Grounding in Narrative. In Steven E. Runge and Christopher J. Fresch (eds.), *The Greek Verb Revisited: A Fresh Approach for Biblical Exegesis*, 163–183. Bellingham, WA: Lexham Press.

— 2016b. Functions of the Copula-Participle Combinations ("Periphrastics"). In Steven E. Runge and Christopher J. Fresch (eds.), *The Greek Verb Revisited: A Fresh Approach to Biblical Exegesis*, 307–326. Bellingham, WA: Lexham Press.

Malone, Andrew S. 2020. Cautious about Conatives. *Novum Testamentum* 62(3). 302–321. https://doi.org/10.1163/15685365-12341668

Nardi, Edoardo. 2025. The Periphrasis 'εἶναι + participle' in Post-Classical Greek: Between Variation and Contact. *Archivio glottologico italiano* 110(1). 33–68.

Nardi, Edoardo, and Domenica Romagno. 2025. Voice, Actionality, and Argument Structure of 'Εἶναι + Participle' Periphrases in Post-Classical Greek. In Klaas Bentein, Eleonora Cattafi and Ezra la Roi (eds.), *Subordination and Insubordination in Post-Classical Greek: From Syntax to Context*, 43–64. Berlin: De Gruyter. https://doi.org/10.1515/9783111180366-003

Ortner, Brian Raff. 2019. The Grammaticalization of Time in the Imperfect Tense in the Greek of the New Testament. PhD thesis (Dallas Theological Seminary), https://www.proquest.com/docview/2288850155?pq-origsite=primo&accountid=13042

Porter, Stanley E. 1992. *Idioms of the Greek New Testament*. Sheffield: JSOT Press.

Probert, Philomen. 2018. Review of Beintein (K.) Verbal Periphrasis in Ancient Greek. Have- and Be- Constructions. *The Classical Review* 68(1). 1–3. https://doi.org/10.1017/S0009840X17001524

Reilig, J. 1965. The Use and Translation of "Kai Egeneto", And it Happened, in the New Testament. *The Bible Translator* 16(4). 153–163. https://doi.org/10.1177/000608446501600401

Rijksbaron, Albert. 2006. *The Syntax and Semantics of the Verb in Classical Greek: An Introduction*. 3rd edn. Chicago, IL: University of Chicago Press.

Robertson, Archibald T. 1914. *A Grammar of the Greek New Testament in the Light of Historical Research*. New York: Hodder & Stoughton.

Rosén, Haiim B. 1957. Die "zweiten" Tempora des Griechischen: zum Pradikatsausdruck beim griechischen Verbum. *Musum Helveticum* 14(3). 133–154.

Rundgren, Frithiof. 1965. The Synoptic Gospels as Language. *Biblica* 46(4). 465–469.

Ryan, Cressida. 2024. Teaching New Testament Greek: What, and How? In Chiara Monaco, Robert Machado, and Eleni Bozia (eds.), *Beyond the Standards: Attic, Koine, and Atticism*, 313–334. Leiden: Brill. https://doi.org/10.1163/9789004687318_013

Siebenthal, Heinrich von. 2019. *Ancient Greek Grammar for the Study of the New Testament*. Bern: Peter Lang. https://doi.org/10.3726/b16377

Smyth, Herbert Weir, and Gordon M. Messing. 2013. *Greek Grammar*. Mansfield Centre, CT: Martino Publishing.

Thrall, Margaret E. 1962. *Greek Particles in the New Testament: Linguistic and Exegetical Studies*. Leiden: Brill.

Toit, Herman C. du. 2017. The Function of the Imperfect Tense-Form in the Narrative Discourse of John's Gospel. *Neotestamentica* 51(2). 209–234. https://doi.org/10.1353/neo.2017.0012

Wallace, Daniel B. 1996. *Greek Grammar Beyond the Basics: An Exegetical Syntax of the New Testament with Scripture, Subject, and Greek Word Indexes*. Grand Rapids, MI: Zondervan.

9. Support-verb constructions with causative force in the Vulgate and Septuagint and their Hebrew and Greek originals (3rd c. BCE to 2nd c. CE)

Roland Hoffmann[1]

1. Introduction

Support-verb constructions consist of a fixed combination of a semantically faded verb, like *dare* (“to give”) and *facere* (“to make”) or *venire* (“to come”) in Latin, which can be called a *support verb* and an *abstract noun* that refers to an entire state of affairs, such as *spes* (“hope”) to *sperare* (“to hope”) and *cogitatio* (“thought”) to *cogitare* (“to think”) in (1) and (2). In Latin linguistics the term “Support-verb constructions” (cf. Pinkster 2015: 74, n.4) is normally used, whereas in German linguistics the term “Funktionsverbgefüge” has been coined and the typological term is “light-verb constructions” (cf. Hoffmann 2023). In the following, the first term, namely “support-verb constructions” (SVCs henceforth), will be used. SVCs consist of a verbal component, the *support verb* (SV henceforth), and a nominal component derived from a noun referring to a state of affairs, the *abstract noun* (AN henceforth). The underlying verbs as *sperare* (“to hope”) in (1) and *cogitare* (“to think”) in (2) are called the *base verbs*. Such SVCs occur in many languages, including Latin and Ancient Greek in all their phases, as well as in German, English, and French. They can mark

1 As the careful reader will easily see, this article is a first approximation to a topic which needs further research as to the corpus material and to the types of support verb constructions. The dataset is accessible here: https://doi.org/10.5287/ora-0zvrz5vnk

©2026 Roland Hoffmann, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.09

Aktionsarten such as inchoative (*in spem venire* "to begin to have a good hope", see (1)) or durative (*in cogitatione versari* "to remain in careful thought", see (2)):

(1) Caesar, *Gallic War*, 1.42.3

[…] magnamque in spem ueniebat [sc. Caesarem] pro suis tantis populique Romani in eum [sc. Ariovistum] beneficiis cognitis suis postulatis fore, uti pertinacia desisteret, dies conloquio dictus est ex eo die quintus.

"Moreover, he [sc. Caesar] began to have a good hope that, in consideration of the signal benefits conferred upon him by Caesar and the Roman people, Ariovistus would abandon his obstinacy when he knew Caesar's demands. A day–the fifth after that–was appointed for the parley." (transl. by Edwards 1917)[2]

(2) Cicero, *Republic*, 1.35

Tum ille. Non possum (equidem) dicere me ulla in cogitatione acrius aut diligentius solere uersari quam in ista ipsa, quae mihi, Laeli, a te proponitur.

"Scipio. I cannot, indeed, assert that any other subject claims more of my interest and careful thought, Laelius, than the one which you now assign to me." (transl. by Keyes 1928)

However, an SVC can also be used to form diatheses such as the passive in (3) (cf. Helbig and Buscha 1989: 81f.) and the causative in (4) (cf. Helbig and Buscha 1989: 83). This causes either a reduction of or an increase in valency by one argument each, as the two instances of Modern High German show:

(3) Das neue Drama von Gerlind Reinshagen kam erstmals in Stuttgart zur Aufführung. = Das neue Drama von G.R. wurde erstmals in Stuttgart aufgeführt.

"Gerlind Reinshagen's new drama premiered in Stuttgart" (own example)

(4) Ich bringe sie zum Lachen.

"I'll make her laugh." (own example)

In the following, only the causative SVCs, which are a type of periphrastic causative forms, will be examined in more detail below with special emphasis on the Latin Bible of the Vulgate and with an outlook to the Hebrew original and the Greek translation of the Septuagint. Since both the Septuagint and the Vulgate are translated texts, all those SVCs where there is a morphological causative in the source language are *per definitionem* causative SVCs apart from the valency increase. This chapter, being descriptive in nature, demonstrates, on the one hand, the importance of SVCs for the expression of causativity and, on the other hand, the increase of this periphrastic means of expression in Latin compared to Ancient Greek.

2 The translations of instances from the classical period of the Latin language are cited according to the Loeb edition.

2. Defining the causative

A causative is the formulation of a state of affairs with an agent as a subject, which does not itself determine the state of affairs, but functions as a caused constituent, a *causee*. This function which is syntactically and semantically defined is brought to carry out the action of the situation by an added causer (Comrie 1989; Dixon 2000). Often, the causer takes the nominative case, while the causee takes an oblique case, for instance a dative case.

This syntagm is therefore a grammaticalised phenomenon, and not only, as van Pottelberge (2007: 436) assumes, one of the stylistic forms that are located "(i)n einem umstrittenen Grenzbereich zwischen Syntax und Phraseologie [in a controversial border area between syntax and phraseology] […] Verb-Substantiv-Verbindungen [verb-noun combinations] […]". Nevertheless, this author names as the four main characteristics (van Pottelberge 2007: 437):

1. A "meaningless" or a "stereotypical everyday verb";
2. a noun that is a verbal abstract and describes a fact;
3. noun and verb form a close and semantic unity, whereby the noun can almost never be omitted;
4. this verb-noun connection can be paraphrased using a synonymous verb.

There are other criteria that vary from author to author (Helbig 1984; Hoffmann 2018). An SVC is the result of a grammaticalisation process. Regardless of van Pottelberge's scepticism, in this chapter such verb-noun syntagms are viewed as grammatical phenomena.

A test for defining a causative is that of *machen* "to make" in languages like German, which form an analytic causative with the meaningless verb *machen* "to make". The German example cited in (4) in Section 1 would allow the following two synonymous clauses according to this *machen* "to make" test.

(5a) *Ich bringe sie zum Lachen.* ("I make her laugh") [repeated from above, (4)]

(5b) *Ich mache, dass sie lacht*. (literally "I make that she laughs.")

3. Causative support-verb constructions in the corpus of the Bible

As is well known, the Bible was written partly in Hebrew (in the Old Testament (OT) with the exception of the Apocrypha written in Greek and a few passages written in Aramaic),[3] partly in Greek (in the New Testament (NT) together with the Apocrypha of the Old Testament). Both parts were translated early.

3 The following passages in the Old Testament are written in Aramaic. Daniel 2:4b-7:28; Esra 4:8-6:18 and 7:12-7:26; Jeremiah 10:11; Genesis 31:47. Cf. Mayer 1965: 35.

This corpus is therefore particularly suitable for a comparison of these three languages, assuming that Biblical Hebrew also has SVCs. Since we have original literature on the one hand with a morphological causative and translation texts partly in Greek and completely in Latin on the other, we can examine a syntagm like the (causative) SVC in several languages at the same time in the Hebrew and Greek original and in the Greek and Latin translation. A causative SVC in the target languages can only be where a morphological causative is in the source language. The parallelism between Hebrew and Latin SVC is at the same time translational compensation and structural and functional equivalence. The definition excludes all examples where there is not an abstract noun (or something similar) and a verb with a general meaning.

An article by Kedar-Kopfstein (1973) is a representative example of previous research. This author examined "the rendering of the Hebrew causative case in the Vulgate." He distinguished between synthetic and analytical rendering of a Hifil form.[4] Nominal examples were cited among the "translations of the root" in addition to the infinitive and the perfect passive participle, namely *peccatum* "sin", *regem* "king", *potum* "drinking", and *fiduciam* "confidence". Without the nomen agentis *regem* "king", the other three, namely *ad peccatum ducere* ("to seduce someone to sin", Nehemiah 13:26), *potum dare* ("to give someone to drink", e.g. Genesis 24:18) and *fiduciam tribuere* ("to inspire someone with confidence", 4 Kings 18:30)[5] as well as two other SVCs, *laborem praebere* ("to cause someone trouble", Isaiah 43:23)[6] and *contumeliis afficere* ("to slander someone", Esra 22:7; quoted in Kedar-Kopfstein 1973: 214, n. 36)—all these five syntagms are indeed causative SVCs.

With regard to Hebrew lexical phrases, we begin with a striking example, namely Sarah's laughter in Genesis 21:6, (6a).

(6a) Genesis 21:6 (BSV)

dixitque Sarra. ***risum fecit*** *mihi Deus. quicumque audierit conridebit mihi.*

"But Sarah said, 'God has **made** me **laugh**; everyone who hears about this will laugh with me.'"

In the Latin bible, there is an SVC with *facere* "to do" and a verbal AN in the object position: *risum fecit* "to make laugh". In addition, there is an explicit causee in the form of a dative next to the SVC. This evidence will be examined in more detail below in (20) with regard to its two models.

We continue with three other favourable examples in (17) to (19), namely with an example that has two different renderings of the Hebrew causative, and because of its familiarity with the sixth petition of the Lord's Prayer (Matthew

4 The Hifil stem serves as a causative in many situations.
5 The non-causative *fiduciam habere* "to have confidence (in)" is more commonly used with this noun (Deuteronomy 28:52; 4 Kings 18:24; Psalms (H) 48:7, etc.).
6 Isaiah 43:23 f. Similarly Acts 17:31: *fidem praebens omnibus* "to grant faith to all".

6:13 par. Luke 11:4) and the related so-called Gethsemane verse (Matthew 26:41 par. Mark 14:38 par. Luke 22:46). These verses, since they are part of the New Testament, only allow comparison with the Greek. Some authors, such as Jeremias (1973) and Schwarz (1987) in particular, assume an Aramaic original version for the so-called sixth petition of the Lord's Prayer (Matthew 6:13 par. Luke 11:4) and have attempted to reconstruct the wording.

(7) Causative *to hope* in a psalm:

(a) Psalms (H) 119:49 (BH)

zeḵōr dāḇār laʿaḇdeḵāḇ ʿal ăšer ***yiḥaltānî***

"Remember your word to your servant, for **you have given me hope**." (NIV)[7]

(b) Psalms 118:49 (LXX)

ζαι μνήσθητι τὸν λόγον σου τῷ δούλῳ σου ᾧ ἐπήλπισάς με

zai mnēsthēti ton logon sou tō doulō sou hō ***epēlpisas*** *me*

"Remember your word to your slave, by which **you buoyed** me **with hope**." (NETS, Pietersma)[8]

(c) Psalms (G) 118.49 (BSV)

[Zain.] *Memor esto uerbi tui seruo tuo, in quo mihi* ***spem dedisti.***

"I will remember your word, your servant, in whom **you have given** me **hope**."

(d) Psalms (H) 118.49 (BSV)

[Zain.] *Memor esto uerbi tui seruo tuo, quo mihi* ***sperare fecisti.***

"Remember the words you spoke to your servant, which **gave** me **hope**!"

(8) Causative *to be tempted* in the sixth intercession of the Lord's Prayer:

(a) Matthew 6:13 (NIV)

Καὶ μὴ **εἰσενέγκῃς** ἡμᾶς **εἰς πειρασμόν**.

kai mē ***eisenegkēs*** *hēmas* ***eis peirasmon.***

"And **lead** us not **into temptation.**"

(b) Luke 11:4 (BSV)[9]

7 The translation of the original examples of the Old and New Testament is taken from the New International Version (NIV).

8 The English translations of the Septuagint instances are taken from Online Version of the "New English Translation of the Septuagint (NETS)", with the translator indicated after the abbreviation, e.g. NETS, Pietersma.

9 "Ein aramäischer Kausativ, der hinter μὴ εἰςενέγκῃς [*mè eisenenkēis*] stehen dürfte, kann sowohl aktives wie permissives Zulassen meinen. So ist es vielleicht möglich, von einem aramäischen Urtext aus den Gedanken einer alleinigen Verantwortung Gottes fernzuhalten" (An Aramaic causative, which is likely the underlying term in μὴ εἰςενέγκῃς [*mè eisenenkēis*] can mean both active and permissive acceptance. Thus, it might be possible to avoid, based on an Aramaic text, the idea of God's sole responsibility). In this exonerating sense Jenni (1992) was arguing: "Aber der griechische Übersetzer, der in seiner Übersetzung von einem aktiven Handeln zu sprechen scheint, hat offenbar solche Probleme nicht" (But the Greek translator, who in his translation seems to speak of active action, apparently does not have

*Et ne nos **inducas in tentationem**.*

“And **lead** us not **into temptation**!”

(9) Causative *to be attempted* in the Gethsemane Verse:

(a1) Matthew 26:41 (NIV)

γρηγορεῖτε καὶ προσεύχεσθε, ἵνα μὴ **εἰσέλθητε εἰς πειρασμόν**.

*grēgoreite kai proseukhesthe, hina mē **eiselthēte eis peirasmon***

“Watch and pray so that you **will** not **fall into temptation**.”

(a2) Luke 22:40 (NIV)

γενόμενος δὲ ἐπὶ τοῦ τόπου εἶπεν αὐτοῖς: προσεύχεσθε, μὴ **εἰσελθεῖν εἰς πειρασμόν**. (NT *Graecae*)

*genomenos de epi tou topou eipen autois: proseukhesthe, mē **eiselthein eis peirasmon**.*

“On reaching the place, he said to them, ‘Pray that you **will** not **fall into temptation**.’”

(a3) Mark 14:38 (NIV)

γρηγορεῖτε καὶ προσεύχεσθε, ἵνα μὴ **εἰσέλθητε εἰς πειρασμόν**. (NT *Graecae*)

*grēgoreite kai proseukhesthe, hina mē **eiselthēte eis peirasmon***

“Watch and pray so that you **will** not **fall into temptation**.”

(b1) Mark 14:38 (BSV)

*Vigilate et orate ut non **intretis in temptationem**.*

“Watch and pray that you **may** not **fall into temptation**!”

(b2) Matthew 26:41 (BSV)

*Vigilate et orate ut non **intretis in tentationem**.*

“Watch and pray, lest you **enter into temptation**!”

(b3) Luke 22:46 (BSV)

*Et ait illis. Quid domitis? Surgite, orate, ne **intretis in tentationem**.*

“And he [sc. Jesus] said to them [sc. his disciples]: ‘Why are you sleeping? Get up! Pray that you **may** not **fall into temptation**!’”

All three examples are undoubtedly causative SVCs. This is demonstrated by a simple paraphrase with *machen* “to make”, which for (7) reads “make me hope”, for (8) “and make that we are not tempted”, and for (9) “make that you are not attempted.” Furthermore, in all three texts the base verbs are transitives (*to hope*

such problems.). Luz 2002: 454, n. 130: “[...] Der Teufel versucht Jesus, aber der Geist versetzt ihn in die Wüste” (The devil tempts Jesus, but the Spirit takes him in to the desert). Luz (2002: 454 n. 129) speaks in a quotation by J. Jeremias of a “permissive nuance of the causative”, which is said to be here. See further also Schwarz (1987).

something, to try someone/something) with a patient consisting of a person. At the same time, there is an argument in the dative case that forms the causee (*mihi* "to me", *nos* "us"). Aside from that, in the same verse (a) we have a Hifil form (*yiḥaltānî*) as the Hebrew equivalent. While the Septuagint in (7b) chooses a verb with the suffix -ίζω (*-ízō*), namely ἐπελπίζω (*epelpízō*) "to pin one's hope in", there are two different versions of this book in the Latin of the Vulgate. In the Hebrew version of Psalms 119:49 in (7d), a rendering of the causative case frequently encountered in late Latin is chosen, namely *facere* "to make" with the infinitive (Chamberlain 1986).[10] In contrast, the Greek version, which dates back to the Septuagint, uses an SVC consisting of *spem* "hope" and *dare* "to give" in (7c).

In (8) we find a word of Jesus from the Sermon on the Mount, namely the sixth petition of the Lord's Prayer in Matthew 6:13 par. Luke 11:4: "Lead us not into temptation." Following Jeremias (1971), it can be assumed that the Greek version was originally based on an Aramaic version. In the back translation, which Jeremias (1973: 92) himself attempted, the SVC *in Versuchung führen* ("to lead into temptation") is the translation of a causative construction present in Aramaic.

In addition, there are other examples in the Vulgate with the same SVC *inducere aliquem in aliquod* "to lead somebody into something" in two apocryphal books, namely in (10) and (11):

(10) Wisdom 15:4 (BSV)

(a) *non enim* ***in errorem induxit*** *nos hominum malae artis excogitatio.*

"For the invention of bad art by humans has not **led** us **into error**."

(b) οὔτε γὰρ **ἐπλάνησεν** ἡμᾶς ἀνθρώπων κακότεχνος ἐπίνοια.

oute gar ***eplanēsen*** *hēmas anthrōpōn kakotekhnos epinoia*

"For the invention of bad art by humans has not **led** us **into error**."

(11) Sirach 4:19 (BSV)

(a) ***timorem et metum et adprobationem inducet*** *super illum [sc. Sapientia filium sapientiae].*

"She [sc. the Wisdom] **will bring fear and anxiety and trials** upon him."

(b) φόβον καὶ δειλίαν ἐπάξει ἐπ᾽ αὐτὸν.

phobon kai deilian epaxei ep' auton.

"She [sc. the Wisdom] **will bring fear and anxiety and trials** upon him."

10 In the entire Vulgate, including the two versions of the Psalms, G and H, there are 239 such instances with *facere* "to make" in the Old Testament alone and 10 in the New Testament (due to the Greek language, which has no genuine causative), making a total of 249 instances. These are the following instance, for example, in the Pentateuch: Genesis 17:6; 24:11; 28:3; 35:11; 41:43; 41:51; 41:52; 47:6; 50:23; Exodus 9:20; 10:19; 17:3; 23:33; Leviticus 1:15; 4:3; 5:9; 6:11; 16:7; Numeri 16:10; 20:5; 31:16; 32:5; Deuteronomy 2:21; 4:36; 13:5; 20:8; 24:4; 27:18: 28:11; 30:9; 32:26; 32:39.

In (10), the causative should be understood to mean "the [...] invention did not make us/cause us to err." The Septuagint uses the verb πλανάω (*planáō*) "to cause to wander" for this, thus having a single verb functioning from a synthetic causative form. A similar synthetic causative form is found in Psalm 118/119, respectively, verse 49, in the form in (39) of ἐπελπίζω (*epelpízō*) "to pin one's hope in."

That both cases are examples of causative SVCs can be recognised by the following criteria. In (10a), the underlying Hebrew verb is a Hifil form, and the Greek version parallels the Hebrew version of a typical causative form with *facere* "to make" and infinitive (*nos sperare fecit* "to make us hope"). In (11), we have a prepositional SVC with *inducere* "to lead in(to)" as an SV and as the AN the three-part accusative object *timorem et metum et adprobationem* "fear, anguish, and trial" (Sirach 4:19).

4. The material of causative support-verb constructions in the Hebrew, Greek, and Latin Bible

4.1 Limits of the test of German *lassen* "to let"

When tracking down further such syntagms, one can use a language such as German, which expresses these constructions periphrastically with *lassen* "to let" and the accusative with infinitive, for example, *man ließ die Kinder kommen* "one let the children come" (Ballweg 1997: 1117.1411.), as suggested by Dixon (2000: 41).[11] Therefore, in the following, all passages with causative *lassen*[12] "to let" (according to the interlinear version in volume 1 by Steurer 1989) are analysed. There are about 250 examples (to be precise, 247). However, the comparison with the Vulgate shows that only a few SVCs can be obtained in this way, namely only the following eight examples (12) to (19), i.e., a total of 8 examples or 3.2%. In the majority of these 247 examples, other renderings are chosen, namely transitive or intransitive verbs or the causative relationship is completely dissolved. In all these cases, causativity ultimately remains unmarked, or the aforementioned *machen* "to make" with an infinitive clause is

11 Compare to the so-called causative complex according to §680 of the Duden, which consists of the auxiliary *lassen* "to let" with the infinitive of a full verb, such as *man* "one" + *lassen* "to let"/*man ließ die Kinder kommen*, literally "one lets the children come". This should be distinguished from the German *sich lassen* "to let oneself" + infinitive, which is a modal passive variant, as can be seen in the following note.

12 German *sich lassen* "to let oneself" can also be used as a modal passive, e.g. *Das Buch läßt sich gut verkaufen = Das Buch kann gut verkauft werden* "the book sells well" (cf. Helbig and Buscha 1989: 187). Thus, *lassen* "to let a state of affairs be done" alone needs not be the right verb, but it must be causative *lassen* "to let".

chosen (Hoffmann 2008, 2014: 164–166).[13]

The eight remaining, (12) to (19), examples are printed here with the Hebrew original in transcription, in which the corresponding causative verb form is analysed.

(12) Genesis 24:18 (NIV)

quae (puella) respondit. "bibe domine mi" celeriterque deposuit hydriam super ulnam suam et ***dedit*** *ei* ***potum****.*

[…] καὶ ἐπότισεν αὐτόν (LXX: no SVC)

kai epotisen auton

wattašᵊqehu	*'al–yɔdɔh*	*kaddɔh*	*wattorɛd*	*wattᵊmaher*	*'ᵃdoniʸ*	*šᵊteʰ*	*watto'mɛr* (BH)
nd-3. SG.F.IMPF.HIFIL+ 3.SG. of *šɔqah* "to-drink": "she-gave-him-water-to-drink"	on-her-hand	(her)-jug	and-she-lowered	and-made-quick	my-lord	drink	and-she [sc. Rebekha] said

"'Drink, my lord,' she said, and quickly lowered the jar to her hands and gave him a drink."

(13) Exodus 10:13 (NIV)

extendit Moses uirgam super terram Aegypti et Dominus ***induxit uentum*** *urentem tota illa die ac nocte* (V)

καὶ ἐπῆρεν Μωυσῆς τὴν ῥάβδον εἰς τὸν οὐρανόν, καὶ κύριος ἐπήγαγεν ἄνεμον νότον ἐπὶ τὴν γὴν ὅλην τὴν ἡμέραν καὶ ὅλην τὴν νύκτα. (LXX)

kai epēren Mōusēs tēn hrabdon eis ton ouranon, kai kurios epēgagen anemon noton epi tēn gēn holēn tēn hēmeran kai holēn tēn nukta.

qɔdiʸm ruaḥ	*nihag*	*waYhwɔʰ*	*Misᵊrayim*	*'al'ɔrɛṣ*	*'ɛt-mattehuʷ*	*Mošeʰ*	*wayyeṭ*
Eastwind	3.SG.M.PIEL of *nɔhag* "to-lead-to": "he-has-led-to"	And Yahweh	Egypt	over-the-land	his-staff	Moses	he-stretch

13 For the number of causative constructions in the five books of Moses with *facere* "to make" and infinitive, see above n. 10 (cf. Hoffmann 2008; 2014: 150–154).

[...] wᵃkɔl=hallɔyᵃlɔʰ	*hahuʷ'*	*kɔl=hayyoʷm*	*bɔ'ɔrɛṣ* (BH)
and-the-whole-night	the same	the-whole-day	over-the-land

"So Moses stretched out his staff over Egypt, and the Lord made an east wind blow across the land all that day and all that night."

(14) Numbers 11:18 (NIV)

*ego enim audiui uos dicere. "Quis **dabit** nobis **escas** carnium? Bene nobis erat in Aegypto."*

καὶ φάγεσθε κρέα [...] λέγοντες Τίς ἡμᾶς ψωμιεῖ κρέα; (LXX)

kai phagesthe krea [...] legontes Tis hēmas psōmiei krea?

bɔśɔr	*yaᵓᵃkilenuʷ*	*miʸ*	*le'mor*	*Yᵃhwɔʰ*	*bə'ɔzᵃney*	*bᵃkiʸtɛm*	*kiʸtoʷb [...]*
meat	3.SG.M.JUSS.HIFIL +1.COM of *'ɔkal*. "to-eat": "he-will-let-us-eat"	who	saying	YHWHs	in-the-ears	you-have-wept	the-good

bᵃMiṣᵃrɔyim	*lɔnuʷ*	*kiʸ-toʷb*
in-Egypt	us/we-were	then

"The Lord heard you when you wailed, 'If only we had meat to eat! We were better off in Egypt!' Now the Lord will give you meat, and you will eat it."

(15) Numbers 16:14 (NIV)

*re uera induxisti nos in terram quae fluit riuis lactis et mellis et **dedisti** nobis **possessiones** agrorum et uinearum an et oculos nostros uis eruere non uenimus*

εἰ καὶ εἰς γῆν ῥέουσαν γάλα καὶ μέλι εἰσήγαγες ἡμᾶς καὶ **ἔδωκας** ἡμῖν **κλῆρον** ἀγροῦ καὶ ἀμπελῶνας, τοὺς ὀφθαλμοὺς τῶν ἀνθρώπων ἐκείνων ἂν ἐξέκοψας. οὐκ ἀναβαίνομεν.

*ei kai eis gēn hreousan gala kai meli eisēgages hēmas kai **edōkas** hēmin **klēron** agrou kai ampelōnas, tous ophthalmous tōn anthrōpōn ekeinōn an exekopsas. ouk anabainomen.*

hᵃbiʸ'otɔnuʷ	*uʷdᵉbaš*	*hɔlɔb*	*zɔbat*	*'ɛl-'ɛrɛṣ*	*lo'*	*'ap (BH)*
2.SG.M.PERF.HIFIL+1.COM.: "you-have-made-us-come"	and-honey	(of)-milk	fluent	in-(a)-land	not	certainly-not

"Moreover, you haven't brought us into a land flowing with milk and honey or given us an inheritance of fields and vineyards. Do you want to treat these men like slaves? No, we will not come!"

(16) Numbers 32:5 (NIV)

*precamurque si inuenimus gratiam coram te, ut **des** nobis famulis tuis eam **in possessionem**, ne facias nos transire Iordanem.*

καὶ ἔλεγον Εἰ εὕρομεν χάριν ἐνώπιόν σου, **δοθήτω** ἡ γῆ αὕτη τοῖς οἰκέταις σου **ἐν κατασχέσει**, καὶ μὴ διαβιβάσῃς ἡμᾶς τὸν Ιορδάνην.

*kai elegon Ei heuromen kharin enōpion sou, **dothētō**, hē gē autē tois oiketais sou **en kataskhesei** kai mē diabibasēs hēmas ton Iordanēn.*

hazzo't	*'ɛt-hɔ'ɔreṣ*	*yuttan*	*bᵊeʸneʸka*	*hen*	*'im-mɔsɔ'nuʷ*	*wayyo'mᵊruʷ*
the-same	ART-land	3.SG.M.IMPF. HOPHAL "it-may-be-given"	in-your-eyes	favour	when-we-have-found	And-they said

laʾᵃaḥuzzɔh	*la-ʿᵃbɔdɛʸkɔ (BH)*
(as)-property	to-your-servants

"If we have found favour in your eyes, they said, let this land be given to your servants as our possession. Do not make us cross the Jordan."

(17) Deuteronomy 3:20 (NIV)

*donec **requiem tribuat** Dominus fratribus uestris sicut uobis tribuit […].* (Vulg.)

ἕως ἂν καταπαύσῃ κύριος ὁ θεὸς ἡμῶν τοὺς ἀδελφοὺς ὑμῶν [...] (LXX)

heōs an katapausē kurios ho theos hēmōn tous adelphous humōn

kɔkɛm	*laʾᵃḥeʸkɛm*	*Yᵊhweh*	*yɔniʸah=ʾᵃšɛr*	*'ad* (BH)
euch-gleich	your-brothers	Yᵊhwɔʰ	3.SG.M.IMPF.HIF. of *nuaḥ*: "to-rest" "he-makes-rest"	until

"[…] until the Lord gives rest to your fellow Israelites as he has to you […]."

(18) Deuteronomy 8:3 (NIV)

*adflixit te penuria et **dedit cibum** manna quem ignorabas tu et patres tui […].* (Vulg.)

καὶ ἐκάκωσέν σε καὶ ἐλιμαγχόνησέν σε καὶ ἐψώμισέ σε τὸ μαννα [...]. (LXX)

kai ekakōsen se kai elimagkhonēsen se kai epsōmise se to manna […]

lo'-yɔdaʿᵊtɔ	*ʾᵃšɛr*	*'et-hammɔn*	*wayyaʾᵊkilᵃkɔ*	*wayyarəʿibɛkɔ*	*wayəʿannəkɔ*
you-did-not-know	which-REL	AKK-ART-manna	3.SG.IMPF.HIFIL of *'ɔkal.* "to-eat" "and-he-made-you-eat"	and-he-let-you-starve	and-he-bent-you

ʾᵃboteʸkɔ	*yadᵊʾuʷn*	*wᵊlo'*
your-fathers	they-had-known	and-not

"He humbled you, causing you to hunger and then feeding you with manna, which neither you nor your ancestors had known, to teach you that man does not live on bread alone but on every word that comes from the mouth of the Lord."

(19) Deuteronomy 28:24 (NIV)

Det *Dominus* ***imbrem*** *terrae tuae puluerem et de caelo descendat super te cinis donec conteraris.*

δῴη κύριος τὸν ὑετὸν τῇ γῇ σου κονιρτόν, καὶ χοῦς ἐκ τοῦ οὐρανοῦ καταβήσεται ἐπὶ σε ἕως ἂν ἐκτρίψῃ σε καὶ ἕως ἂν ἀπολέσῃ σε (LXX)

dōē kurios ton hueton tē gē sou konirton kai khous ek tou ouranou katabēsetai epi se heōs an ektripsē se kai heōs an apolesē se

min-ha-ššɔmayim	*wᵊʿɔpɔr*	*'ɔbɔq*	*'arᵊṣᵊkɔ*	*eṭ-mᵊtar'ɛt*	*Yᵊhwɔʰ*	*yitten*
of-the-heavens	and-dust	quicksand	your-land-DAT	rain	Yahweh	3.SG.M.IMP. Qal of *natan.* "to-give" "he-gives"

"The Lord will turn the rain of your country into dust and powder; it will come down from the skies until you are destroyed."

We look now whether in the Hebrew original and in the Septuagint translation there is already an SVC. While the Vulgate has in all eight examples SVCs, in the Hebrew original there is an SVC in only one case, namely (19), Deuteronomy 28:24: *Yᵊhwɔʰ yitten 'eṭ-mᵊtar'ɛt* "Yahwe makes it rain", which corresponds to the causative construction, and the verb is in the Qal:[14] "Yahwä makes it rain". In five cases the verb is in the Hifil (Genesis 24:18, see (12); Numbers 11:18, see (14); Numbers 16:14, see (15); Deuteronomy 3:20, see (17); Deuteronomy 8:3, see (18)), in one example in the Piel (Exodus 10:13, see (13)), in another in the Hofal

14 The Qal is the basic stem of the Hebrew verb.

(Numbers 32:5, see (16)), which is normally the passive to a Hifil.

As for the Septuagint, half of the forms are single-word forms, three of which have the suffix -ιζ- (*-iz-*): ποτίζω (*potízō*) "to let drink", Genesis 24:18, see (12); 2x ψωμίζω (*psōmízō*) "to feed", Numbers 11:18, see (14), Deuteronomy 3:20, see (17), and Deuteronomy 8:3, see (18). While the *Biblia Hebraica* uses morphological means in the form of verb stems, in the LXX, lexical-derivative means partially fulfills this function, namely verbs with the suffix -ιζ- (*-iz-*). The other four forms are formed from two or more words and can be described as an SVC except for one example, namely Exodus 10:13 (ἐπάγειν ἄνεμον (*epágein ánemon*), see (13)), Numbers 16:14; (διδόναι κλῆρον (*didónai klêron*), see (15)), Deuteronomy 28:24 (διδόναι τὸν ὑετόν (*didónai tòn hyetón*), see (19)). The example in Numbers 32:5, see (16a), is an SVC with the SV as a prepositional phrase.

(16a) Numbers 32:5 (LXX)

[...] **δοθήτω** ἡ γῆ αὕτη τοῖς οἰκέταις σου **ἐν κατασχέσει** [...]

[...] ***dothētō*** *hē gē hautē tois oiketais sou* ***en kataskhesei*** *[...]*

"May this land be given to your families as property"

Thus, from the *Biblia Hebraica* (BH) to the Vulgate, the following can be summarised. The text of the BH has predominantly single-word forms, with only one SVC in (12). The Septuagint has only half single-word forms, and four multi-word forms, three of which are SVCs. The Vulgate, on the other hand, has only multi-word forms in the form of SVCs. However, a further group of examples below will show that this tendency is not found everywhere, especially with regard to the Hebrew *Biblia Hebraica*.

A clear causative SVC is the example already considered above as (14): the SV here is *facere* "to make", the AN is *risum* "laughter", while the nominative *Deus* "god" forms the causer and the dative *mihi* "to me" the causee.[15] What is the Hebrew and Greek text for this?

(20a) Genesis 21:6 (BSV) [repeated from above, (6a) and (12)]

Dixitque Sarra. ***risum fecit*** *mihi Deus. quicumque audierit, conridebit mihi.*

"But Sarah said, 'God has **made** me **laugh**; everyone who hears about this will laugh with me.'"

The two sources for the Vulgate, the text of the *Biblia Hebraica* and the Septuagint version, are:

15 An Itala manuscript has the noun *gaudium* "joy" instead of *risum* "laughter".

(20b)	*watto'mɛr*	*Śɔrɔʰ.*	*sᵊhoq*	*ʿɔsɔh liʸ*	*liʸ*
	3.SG.F.NARR. 'mr. "to-talk", "to-say"	Sarah	"(a-)laughter"	3.SG.M.PERF. ʿsh. "to-do", "to-make"	to me = "me".DAT

-liʸ	*yiṣᵊḥaq*	*kɔl-haššomeaʿ*
-me.DAT	3.SG.IMPF. *ṣᵊhoq*. "to-laugh"	jeder-PTZP-QUAL *'mr.* "to-talk", "to-say"

(20c) Genesis 21:6 (NETS, Hiebert)

εἶπεν δὲ Σαρρα· Γέλωτά μοι ἔποιησεν ὁ κύριος.

eipen de Sarra: Gelōta moi epoiēsen ho kurios

"And Sarra said, 'The Lord has made laughter for me.'"

In all three versions, the SV is a synonymous verb, a so-called common verb or an "Allerweltsverb" (Happ 1976: 264) (*ʿɔsɔh*—ποιεῖν—*facere* "to make") with the AN in the accusative case (*səhoq*—γέλωτα—*risum* "laughter"); thus, it is translated literally in both translations. Furthermore, the causee is in the dative case (*lij*,[16] μοι *mihi* "to me") in all three versions.

4.2 Two other tests: *Verba timendi* and the two most frequent support verbs

The *lassen* "to let" method has yielded few instances, which is more a limitation of the test than an indication of the robustness of the causative SVC category. It had the only function to show causative constructions in the (Pentateuch of the) Old Testament. Thus, another method must be found. Since SVCs are a grammatical means, they are everywhere possible and never translation-induced. We take two methods and combine them. On the one hand, we choose certain nouns that are very productive, and on the other, we choose two other methods of finding SVs. Such typical ANs are verbs of sensory and mental perception, especially those of fear, such as *terror* "panic", *formido* "fearfulness", *metus* "dread", *timor* "fear", but also the already cited *spes* "hope" (Section 4.2.1). The two most common SVs are *dare* "to give" and *facere* "to make" (Section 4.2.2). Methodologically, we therefore choose, on the one

16 It is well known that Biblical Hebrew does not have any case markers. However, the dative case is expressed with the proclitic preposition *lə* and the accusative may be marked by the so-called nota accusativi *'et* (cf. Genesis 1.1).

hand, a semantic and, on the other hand, a more statistical approach.

4.2.1 Abstract nouns with the meaning of FEAR and TERROR

A brief explanation of this specific semantic verb group is required, as they are used here as causative verbs. A look at German grammar is helpful here. The verb *(er)schrecken* "to frighten", which belongs to the AN *Schrecken* "horror", has two meanings. According to Grimm's Dictionary (2004: 970), it is, on the one hand, an intransitive *erschrecken* "to frighten", meaning "to get a fright, to scare"; on the other hand, this verb has the meaning "to put someone in fear and terror." Thus, in the second variant, it is undoubtedly a causative verb with the meaning "to cause someone else to get a fright."[17] Both meanings are reflected in German with different tense formation. While the intransitive verb forms a strong preterite, for example, *when the cat suddenly entered the room, I was frightened* translates as "ich war erschrocken", the causative verb has a weak preterite, for example, *the cat frightened me* (*when it suddenly stood in front of me)* translates as "sie erschreckte mich" or "ich war erschreckt". In Latin, a noun like *terror* "panic", *formido* "fearfulness", or *metus* "dread" derives from a causative verb.[18] A first striking aspect is that the same verb *dare* "to give" is used as the SV in all eight Latin examples. It will be further investigated whether this is a typical feature of the SVC in the Vulgate, and at the same time, this verb will need to be examined more closely in terms of semantics and valency grammar.

Table 9.1 shows a number of examples. In the following, we will first examine the extent to which an SV already exists in Hebrew and which form the Septuagint then chooses in these examples, before further investigating the type of causation according to the syntactic and semantic parameters identified and explained above. The main focus will be on the SV, because by this method it will be examined how the individual biblical languages—that is, Latin, Greek, and Hebrew (or, instead of Hebrew, only Greek in the case of apocryphal texts not preserved in the Hebrew Bible) relate to these parameters. In this way, we obtain a total of 19 examples in (21) to (39).

17 See Helbig and Buscha (1995: 61): "Unter den Tätigkeitsverben ist die Subklasse der kausativen Verben auszusondern [...]. Sie können in *zwei* verschiedene semantische Prädikate (Tätigkeit und Vorgang) zerlegt werden und stellen folglich eine Kombination aus Tätigkeit und Vorgang dar. Er *zerbrach* die Tasse. = Er *verursachte*, dass die Tasse *zerbrach*" (Among the action verbs, the subclass of causative verbs must be distinguished, […]. They can be broken down into two different semantic predicates (ACTION and PROCESS) and consequently represent a combination of action and process. He *broke* the cup = He *caused* the cup *to break*).

18 See *Thesaurus Linguae Latinae*, s.v. *formido* and s.v. *metus*.

Vulgate	Passage	Biblia Hebraica—verb phrase	Septuagint—verb phrase
Formidinem/terrorem mittere "to send fearfulness/panic"			
(21) ***terrorem*** *meum* ***mittam*** *in praecursum tuum et occidam omnem populum ad quem ingredieris.*	Exodus 23:27	SVC. *ʾɛt-ʸmɔtiʸ* *ʾᵃšallaḥ*	τὸν φόβον ἀποστελῶ *ton phobon apostelō*
"I will send my terror ahead of you and kill every nation you invade."			
(22) *hodie incipiam* ***mittere terrorem atque formidinem*** *tuam in populo qui habitant sub omni caelo ut audito nomine tuo paueant et in morem parturientium contremescant et dolore teneantur.*	Deuteronomy 2:25	SVC. *ʾɔḥel tet* *paḥᵊdᵊkɔ* *wᵊyirᵊʾɔtᵊkɔ*	SVC δοῦναι τὸν τρόμον σου καὶ τὸν φόβον σου *dounai ton tromon sou kai ton phobon sou*
"Today I will begin to send terror and horror before you to the nations under the whole heaven, so that when they hear your name, they will be terrified and tremble like a woman in labour, seized with pain."			
(23) *induxi eis super matrem adulescentis uastatorem meridie* ***misi*** *super ciuitates repente* ***terrorem****.*	Jeremiah 15:8	SVC (verb in Hifil) *hippalᵊtiʸ* *ʿɔlɛʸhɔ […]* *ʿiʸr* *uʷbɛhɔloʷt*	SVC ἐπέρριψα ἐπ' αὐτὴν ἐξαίφνης τρόμον καὶ σπουδήν *eperripsa ep' autēn exaiphnēs tromon kai spoudēn*
"[His widows have become more numerous to me than the sand of the sea. I have brought them a desolator against the young man's mother at midday.] I have sent sudden terror upon the cities."			
Formidinem/pavorem/terrorem dare "to give fearfulness/trembling/panic"			
(24) *omnes hii incircumcisi interfectique et cadentes gladio quia* ***dederunt formidinem*** *suam in terra uiuentium.*	Ezechiel 32:26	SVC *nɔtᵊnuʷ* *ḥittiʸtɔm*	οἱ δεδωκότες τὸν φόβον αὐτῶν *hoi dedōkotes ton phobon autōn*
"All these are uncircumcised, and they were killed and fell by the sword, for they spread terror in the land of the living."			

Vulgate	Passage	Biblia Hebraica—verb phrase	Septuagint—verb phrase
(25) ***da*** *illis* ***formidinem*** *et tabefac audaciam uirtutis eorum et commoueantur contritione sua.*	1 Maccabees 4:32		SVC δὸς αὐτοῖς δειλίαν καὶ τῆξον καὶ θράσος ἴσχυος αὐτῶν *dos autois deilian kai tēxon kai thrasos iskhuos*
"Strike terror into them and melt away the pride of their strength, and let them be shaken by destruction."			
(26) *nullus stabit contra uos* ***terrorem*** *uestrum* ***et formidinem dabit*** *Dominus Deus uester super omnem terram quam calcaturi estis sicut locutus est uobis.*	Deuteronomy 11:25	SVC *paḥᵃdᵃkɛm* *uʷ-moʷraʾᵃkɛm* *yitten YHWH*	SVC τὸν τρόμον ὑμων καὶ τὸν φόβον ἐπιθήσει κύριος *ton tromon humōn kai ton phobon epithēsei kurios*
"The Lord your God will strike terror and dread into the whole land you are about to enter, just as he has told you."			
(27) *diuulgatumque est nomen Dauid in uniuersis regionibus et Dominus* ***dedit pauorem*** *eius super omnes gentes*	1 Chronicles 14:17	SVC *wa-YHWH* *nɔtan ʾɛt-* *paḥᵃdoʷ* *ʿal-kol* *haggoʷyim*	Vc καὶ κύριος ἔδωκεν τὸν φόβον αὐτοῦ *kai kurios edōken ton phobon autou*
"[...] and the Lord put fear of him upon all nations."			
inducere aliquem in pavorem "to make somebody tremble"			
(28) *quia haec dicit Dominus ecce ego* ***dabo*** *te* ***in pauorem****, te et omnes amicos tuos.*	Jeremiah 20:4	SVC *hinnᵃniʸ* *notɛnᵃkɔ* *lᵃmɔgoʷr*	SVC ἐγὼ δίδωμί σε εἰς μετοικίαν *egō didōmi se eis metoikian*
"Behold, I will make you and all your friends tremble with fear."			
terrorem inducere "to induce panic"			

Vulgate	Passage	Biblia Hebraica—verb phrase	Septuagint—verb phrase
(29) *ecce ego* ***inducam*** *super te* ***terrorem*** *ait Dominus Deus exercituum ab omnibus qui sunt in circuitu tuo.*	Jeremiah 49(=30):5	SVC *hinnᵃniʸ mebiʸ' ʿɔlayik paḥad*	SVC φόβον φέρω *phobon pherō*
"Behold, I will bring terror upon you, says the Lord, the God of hosts, from all those around you."			
memoriam facere "to make remember"			
(30) ***memoriam fecit*** *mirabilium suorum misericors et miserator Dominus.*	Psalms 110 (G+ H):4	SVC *zekɛr ʿɔśɔʰ lᵃnipᵃlᵃ'otɔʸw*	SVC μνείαν ἐποιήσατο τῶν θαυμασίων αὐτοῦ *mneian epoiēsato tōn thaumasiōn autou*
"He has brought to mind his wonders; the Lord is gracious and merciful."			
(31) *non cesso gratias agens pro uobis* ***memoriam*** *uestri* ***faciens*** *in orationibus meis.*	Ephesians 1:16		οὐ παύομαι εὐχαριστῶν ὑπὲρ ὑμῶν μνείαν ποιούμενος ἐπὶ τῶν προσευχῶν μου. *ou pauomai eukharistōn huper humōn mneian poioumenos epi tōn proseukhōn mou*
"That is why I never cease to give thanks for you, keeping you in my prayers."			
(32) *gratias agimus Deo semper pro omnibus uobis* ***memoriam facientes*** *in orationibus nostris sine intermissione*	1.Thessalonians 1:2		Εὐχαριστοῦμεν τῷ θεῷ πάντοτε περὶ πάντων ὑμῶν μνείαν ποιούμενοι ἐπὶ τῶν προσευχῶν ἡμῶν, *... eukharistoumen tō theō pantote peri pantōn humō mneian*
"We always thank God for all of you and remember you constantly in our prayers."			

Vulgate	Passage	Biblia Hebraica—verb phrase	Septuagint—verb phrase
(33) *Gratias ago Deo meo semper* ***memoriam*** *tui* ***faciens*** *in orationibus meis.*	Philemon 1:4		Εὐχαριστῶ τῷ θεῷ μου πάντοτε μνείαν σου ποιούμενος ἐπὶ τῶν προσευχῶν μου. *eukharistō tō theō mou pantote mneian sou poioumenos epi tōn proseukhōn*
"I always thank my God when I remember you in my prayers."		–	
obliuionem facere "to cause oblivion"			
(34) *malitia horae* ***obliuionem facit*** *luxuriae magnae et in fine hominis denudatio operum illius.*	Sirach 11:29		κάκωσις ὥρας ἐπιλησμονὴν ποιεῖ τρυφῆς, *kakōsis hōras epilēsmonēn poiei*
"The evil of an hour causes the forgetting of great abundance, / and at the end of man stands the exposure of his works."			
gaudium dare "to give joy"			
(35) *timor Domini delectabit cor et* ***dabit laetitiam et gaudium*** *in longitudine dierum.*	Sirach 1:12		SVC φόβος κυρίου [...] δώσει εὐφροσύνην καὶ χαρὰν καὶ μακρο-ημέρευσιν. *phobos kurious [...] dōsei euphrosunēn kai kharan kai makroēmereusin*
"The fear of the Lord will gladden the heart/and bring joy and happiness in a long succession of days."			

Vulgate	Passage	Biblia Hebraica—verb phrase	Septuagint—verb phrase
(36) *religiositas custodiet et iustificabit cor* ***iucunditatem atque gaudium dabit***	Sirach 1:18		στέφανος σοφίας φόβος κυρίου ἀναθάλλων εἰρήνην καὶ ὑγίειαν ἰάσεως. *stephanos ophias phobos kurious anathallōn eirēnēn kai hugieian iaseōs*
"Piety will preserve the heart and make it righteous; it will bring well-being and joy."			
tristitiam dare "to give grief"			
(37) ***tristitiam*** *non* ***des*** *animae tuae et non adfligas temet ipsum in consilio tuo*	Sirach 30:22		SVC Μὴ δῷς εἰς λύπην τὴν ψυχήν σου... *Mē dōs eis lupēn tēn psukhēn sou*
"Do not give sadness to your soul, and do not grieve yourself with your advice!"			
(38) *cor pravum* ***dabit tristitiam*** *et homo peritus resistet illi*	Sirach 36:22		SVC καρδία στρεβλὴ δώσει λύπην... *kardia streblē dōsei lupēn*
"An evil heart will cause sadness, and an experienced man will resist it."			
spem dare "to make hope"			
(39) *Memor esto uerbi tui seruo tuo in quo mihi* ***spem dedisti****.*	Psalms (G) 118:49	One-Word Form Piel *yiḥalᵊtɔniʸ*	One-Word Form ἐπήλπισάς *epēlpisas*
"Remember the words you spoke to your servant, which gave me hope."			

Table 9.1 SVCs of special *verba sentiendi.*

These nineteen examples from (21) to (39) are extremely revealing. A look at the Hebrew and Greek texts shows that only in one and the same place, namely in Psalms 118: 49, synthetic forms are found: a Piel form in the Hebrew original and a transitive verb ending in -ίζω *-izō* in the Septuagint. In all other cases, a similar syntagm is used as in the Latin SVC. The conclusion is obvious that these are also SVCs. This is supported by the fact that some similar verbs are used: four times in the Hebrew, *nɔtan* "to give" (Deuteronomy 11:25; 1 Chronicles 14:17; Jeremiah 20:4; Ezechiel 32:26) occurs as an SV, and once *ʿɔśɔh* "to make" (Psalms 110/111:4). In Greek, διδόναι (*didónai*) "to give" occurs eight times (Deuteronomy 2:25; 1 Chronicles 14:17; Jeremiah 20:4; Esra 32:26; Sirach 1:12; 30:21; 36:20; 1 Maccabees 4:32) as an SV and five times, including once as ποιεῖν (*poieîn*) "to make, to do" and five times as ποιεῖσθαι (*poieîsthai*) "to make, to do" (Sirach 11:27; Psalms 110/111:4; Ephesians 1:16; 1 Thessalonians 1:2; Philemon 1:4). In a Hebrew example in Jeremiah 15:8 in (23), the SV is also in the Hifil; thus, there is a double marking.

The proportion of synonymous SVs *dare*, διδόναι (*didónai*) "to give", and *nɔtan* "to give" is interesting. This proportion is highest in the Vulgate, with 10 examples, and lowest in the *Biblia Hebraica*, with 4 out of 10 evaluable examples. The Septuagint has a higher proportion, with 8 examples. The *facere*-ποιεῖν (*poieîn*) "to make" examples, in contrast, are fewer. They amount to 5 examples in the Vulgate, also 5 in the Septuagint, and 1 in the *Biblia Hebraica*. Table 9.2 shows the occurrence of SVs in the three biblical sources, in the Hebrew original text, in the Septuagint and in the Vulgate or in the Greek version of the *Novum Testamentum Graecae*.

Vulgata	*Biblia Hebraica*	Septuagint
10%	4/10	8%

Table 9.2 *Dare* "to give" in the three corpora.

Table 9.3 makes the findings clear in the Vulgate alone. The SV *dare* "to give" is by far the most common causative verb here.

SN + *dare* "to give"	10x
SN + *facere* "to make"	5x
SN + *mittere* "to send"	3x
SN + *adferre* "to bring to"	1x
SN + *inducere* "to bring upon"	1x

Table 9.3 Support verbs in the Vulgate.

It promises to be insightful to examine how these syntagms appear in the two translations, since the Hebrew corpus is too small. This involves examining, first, the position of ANs and SVs, and second, the question of possible hyperbata. Table 9.4 shows the result of two corpora of this chapter.

		Noun – Verb	Verb – Noun	Without space (not discontinuous)	With space (discontinuous)
Septuagint[19]	Σ = 21	6x	15x	6x	15x
Vulgate[20]	Σ = 38	24x	14x	22x	16x

Table 9.4 Position of support verb and abstract noun.

There is some variation in both parameters. With regard to discontinuity, it should be noted that in some of the examples, for example, in (12), (14) to (16), and (25), a causee stands between the two elements, (40) and (41), but there are also counterexamples in which the causee appears before, (42), or after both elements, (43).

(40) Genesis 24:18 (NIV) [repeated from above, (6) and (12)]

*quae (puella) respondit. "bibe domine mi" celeriterque deposuit hydram super ulnam suam et **dedit** ei **potum**.*

"'Drink, my lord,' she said, and quickly lowered the jar to her hands and gave him a drink."

(41) Numbers 16:14 (NIV) [repeated from above, (15)]

*re uera induxisti nos in terram quae fluit riuis lactis et mellis et **dedisti** nobis **possessiones** agrorum et uinearum an et oculos nostros vis eruere non uenimus.*

"Moreover, you haven't brought us into a land flowing with milk and honey or given us an inheritance of fields and vineyards."

(42) Psalms 118:49 (NIV) [repeated from above, (39)]

*Memor esto verbi tui servo tuo in quo mihi **spem dedisti**.*

"Remember your word to your servant, for you have given me hope."

19 The 21 instances of the Septuagint are: (19a), (21a), (23a), (24a), (27a), (28d), (29) to (36), (38), (42) to (46), (52).

20 The 38 instances of the Vulgate are: (15c), (16), (17b1-3), (18a), (19a), (20) to (27), (28a), (29) to (46), (52)f., (59), (66).

(43) Deuteronomy 28:24 (NIV) [repeated from above (19)]

δῴη κύριος τὸν ὑετὸν **τῇ γῇ** σου κονιορτόν, καὶ χοῦς ἐκ τοῦ οὐρανοῦ καταβήσεται ἐπὶ σέ, ἕως ἂν ἐκτρίψῃ σε καὶ ἕως ἂν ἀπολέσῃ σε.

*dōē kurios ton hueton **tē gē** sou koniorton, kai khous ek tou ouranou katabēsetai epi se, heōs an ektripsē se kai heōs an apolesē se*

"The Lord will turn the rain of your country into dust and powder; it will come down from the skies until you are destroyed."

In the Septuagint, a clear majority of examples have the verb in front, while in the Vulgate, the opposite is the case: the AN is in front in most cases, although not as clearly as in the Septuagint. Discontinuity is more common in the Septuagint than the direct placement of both elements. In the Vulgate, examples without hyperbata slightly predominate. If we assume a development from the language of the Septuagint to the language of the Vulgate, then a tendency toward AN–SV placement and juxtaposition without intervening elements can be observed. Given the scarcity of the material, it is not possible to say more. After all, even the prototypical Hebrew passage in Genesis 21:6 shows AN–SV placement with no discontinuity.

4.2.2 The two most frequent support verbs: dare "to give" and facere "to make"

potestatem dare "to give permission"	35x
hereditatem dare "give an inheritance"	21x (of which perfect participles: 3x)
fructum facere "to bring/yield fruit"	16x
potum dare "to let drink"	13x[21]
timorem dare "to scare"	9x
requiem dare "to give rest"	12x[22]

Table 9.5 The most frequent SVCs.

21 That *potum dare* "to let drink" is a real causative SVC can be seen from the findings that in the source language in all 13 instances (Genesis 24:14; 24:18; 29:7; Exodus 2:19; 32:20; 2 Samuel 23:15; Psalms 77:15; 79:6; Isaiah 43:20; Jeremiah 8:14; 16:7; Habakkuk 2:15) *potum dare* "to let drink" is the translation of a form of Hifil *šāqah* "to let drink". In the Septuagint instances, the Hifil forms are translated with the verb ποτίζω *potizō* "to let drink".

22 In one place, namely in Deuteronomy 3:20, *requiem* "rest" is found with *tribuere* "to assign to": *donec requiem tribuat Dominus fratribus vestris* "until God gives a resting place to your brothers". This verb also occurs with ANs other than those used in SVCs, such as *hereditas* "inheritance" (Numbers 33:54), *pluvia* "rain" (Deuteronomy 28:12), and *fiducia* "trust" (4 Kings 18:30).

In this corpus analysis, we focus on SVs and attempt to fully capture the two most common, namely *facere* "to make" and *dare* "to give". There are 93 instances of *facere* "to make" in the Old Testament and 49 in the New Testament, for a total of 142 occurrences. Of the SV *dare* "to give" there are 285 instances in the Old Testament and 99 in the New Testament, for a total of 384 occurrences. All instances of *facere* "to make" and *dare* "to give" thus constitute 526 occurrences. Among the individual instances, there are some that occur multiple times. The five most common are shown in (44) to (48) below.

4.2.2.1 Criteria of causativity in the case of the *facere* "to do" instances

A first question regarding the SV *facere* "to make" is when, that is to say, under which contextual conditions, *facere*-SVCs function causatively, since these syntagms are usually simply used suppletively, i.e., instead of one-word forms. For example, in (44), *misericordam facere* "to show kindness" is an SVC which functions suppletively, while *oblivionem facere* "to cause the forgetting" in (46) is a causative one. In the case of *dare* "to give", this problem of ambiguity does not exist.

(44) 2. Samuel 2:5 (BSV)

misit [...] David nuntrios ad viros Iabesgalaad et dixit ad eos. "Benedicti vos domino, qui ***fecistis*** *mihi* ***misericordiam*** *hanc cum domino vestro Saul".*

"David sent messengers to the men of Jabesh-Gilead and said to them, Blessed be you by the Lord, who have shown this kindness to your lord Saul and have buried him!"

(45) Isaiah 29:14 (BSV)

ideo ecce ego addam, ut ***admirationem faciam*** *populo huic miraculo grandi et stupendo peribit enim sapientia a sapientibus eius et intellectus prudentium eius abscondetur.*

"Behold, therefore, I will continue to cause this people to marvel by a great and astonishing miracle, for the wisdom of its wise men shall perish, and the understanding of its prudent men shall be hidden."

(46) Sirach 11:29 (BSV)

Malitia horae ***oblivionem facit*** *luxuriae magnae, et in fine hominis denudatio operum illius.*

"The misfortune of an hour causes the forgetting of great abundance, and at the end of man's life lies the exposure of his deeds."

In instances (44) to (46), the question of why the example in (44) is not causative, whereas examples (45) and (46) are causative, is raised. While the SVC in (44) *fecistis* [...] *misericordiam* "you have shown kindness" has a normal transitive function and does not have a causee in the form of a dative complement, the example *admirationem faciam* "I will cause ... to marvel" in Isaiah 29:14 (45) has a dative that functions as a causee, *populo* "people". Furthermore, the subject in the nominative case, namely *malitia horae* "misfortune of an hour" in (46), functions as a causer. A similar example of a causative SVC with *adferre* "to bring" as an SV is, for example, *spem adferre* "to make hope" (39), which occurs 20 times in Cicero's philosophical writings.[23] It usually has an accusative plus infinitive construction as a second complement and a dative as the causee. So, the answer is that transitive verbs that do not have constituents acting as causee or causer cannot, however, function causatively, but remain transitive even as an SVC (Hoffmann 2005: 555).

Here, the difference between transitive and causative must be discussed. The first term, transitivity, is also used as a generic term, of which the term causativity forms a subclass. However, the term transitive is also used in a narrower sense when it means non-causative. Morphologically, a causative can be marked in various ways, including the form of an SVC. Overlaps between the two phenomena only occur in the case of verbs with an accusative object. Here, the subject of a transitive verb differs in role from the subject of a causative verb form. In the transitive case, the subject is directly involved in the execution of the action, i.e., it is the constituent that acts as the agent and triggers or initiates the action and also executes it. In the causative case, on the other hand, the subject is the causer, i.e., it merely initiates the action, while another constituent executes it. The forms are the same in both cases, except that in the causative form, a causee usually appears in a different case. However, similar to the agent in the passive voice, a causee may also be missing. In this case, it can be recognised pragmatically whether the action is initiated by the subject or also performed at the same time. In the first case, the subject is often also an AN (*emissio* "release" in (47) and *exspectatio* "expectation" in (48)) (cf. Hoffmann 1996: 209f.), because causativity is strictly speaking the juxtaposition of two states of affairs.

23 Instances for *adferre* "to bring" + AN in Cicero's philosophical writings are: Cicero, *On Divination*, 2.62 (*emissio feminae anguis mortem adferebat Ti. Graccho* "the release of the female snake was to be fatal to Tiberius Gracchus"); Cicero, *On the ends of goods and evils*, 1.53 (*quia [sc. iustitia] iucunditatis vel plurimum afferat* "because (sc. justice) is so highly productive for itself"); Cicero, *On the ends of goods and evils*, 2.10 (*dulcem motum afferant sensibus* "they cause a pleasant activity to the senses"); Cicero, *Tusculan Disputations*, 1.86 (*cum M. Antonius summam spem salutis bonis omnibus attulisset* "when Marc Anthony had inspired all good patriots with high and beneficent hopes"): cf. Hoffmann 1996: 209.

(47) Cicero, *On Divination*, 2.62

*Ego tamen miror, si **emissio** feminae anguis mortem adferebat Ti. Graccho, emissio autem maris anguis erat mortifera Corneliae, cur alteram utram emiserit [...]*

"And yet what surprises me is this: If the release of the female snake was to be fatal to Tiberius Gracchus and that of the male was to be the death of Cornelia, why in the world did he let either snake escape?" (transl. by W. A. Falconer)

(48) Cicero, *Philippics*, 5.31

*quorum [legatorum] **exspectatio** dubitationem belli adfert.*

"To wait for them [sc. the envoys] will raise doubts about the war [...]." (transl. by D. R. Shackleton Bailey)

Since most SVCs are formed with *dare* "to give" and this verb is trivalent, in a further study it would be necessary to analyse the extent to which there is a third complement, which, as is well known, is a causee in causative function.

4.3 Corpus analysis of all causative support-verb constructions with *facere* "to make" and *dare* "to give"

To this end, we are conducting a corpus analysis of the entire Bible, focussing on the two most common verbs, namely *facere* "to make" and *dare* "to give". Using the Vulgate as our starting point we only collect SVCs with a causative meaning. In a further step, we examine what is written in the Greek version of the Septuagint and in the original Hebrew version of the *Biblia Hebraica*.

There is a total of 526 examples. Of these, 374 are causative SVCs. That is exactly 71.1%. In the following two Tables 6 and 7, all the examples are compiled in alphabetical order according to the ANs.

Σ = 20:
fulgura "lightning" (2x), *laetitia "joy"*, *memoria* "remembrance" (2x), *misericordia* "compassion", *oblivio* "oblivion", *pavor* "dread" (3x), *pluvia* "rain" (3x), *remissio* "returning", *risus* "laughter", *salus* "health" (2x), *silentium* "silence" (2x), *stupor* "stupor"

Table 9.6 All SVCs in the Bible corpus for the verb *facere* "to make".

Σ = 394:

adflictio “pain” (3x), *aedificatio* “building”, *benedictio* “praising” (2x), *benignitas* “kindness”, *captio* “fraud”, *cibus* “food”, *cinis* “ashes”, *circumcisio* “cutting”, *claritas* “brightness” (2x), *commentio* “mention”, *commotio* “excitement” (2x), *commutatio* “change” (2x), *conceptus* “conception”, *concupiscientia* “arousal” (2x), *consolation* “comfort”, *constantia* “perseverance”, *corruptio* “corruption”, *cura* “care”, *custodia* “protection”, *desertum* “solitude”, *desiderium* “longing”, *despectio* “looking down upon”, *dispensatio* “management”, *dolor* “pain”, *fides* “trust”, *fiducia* “confidence”, *flos* “flower”, *fluctuatio* “wavering” (2x), *formido* “fearfulness” (4x), *fortitudo* “strength” (4x), *fructus* “enjoyment” (15x), *gaudium* “joy” (5x), *gloria* “glory” (15x), *grando* “hail”, *gratia* “favour” (16x), *hereditas* “inheritance” (23x), *honos* “honour” (5x), *ignis* “fire” (4x), *imber* “rain”, *imperium* “power” (3x), *indicium*/*-a* “notice” (2x), *inspiratio* “inspiration”, *intellectus* “discernment” (9x), *intellegentia* “intelligence” (3x), *iudicium* “judgment”, *iustificatio* “justification” (2x), *laetitia* “exultation” (5x), *laus* “praise”, *lucerna* “lamp”, *luctus* “grief”, *lumen* “light” (3x), *lux* “light”, *magnificentia* “nobleness” (2x), *maledictio* “abuse” (2x), *memoria* “remembrance”, *mensura* “measurement”, *nix* “snow” (2x), *oblivio* “oblivion” (2x), *obproprium* “reproach” (2x), *odor* “smell”, *pabulum*/*a* “nourishment” (2x), *paenitentia* “repentance” (3x), *pars* “part”, *pavor* “trembling” (2x), *pax* “peace” (3x), *perditio* “ruin”, *petitio* “request” (4x), *placatio* “quieting”, *pluvia*(e) “rain” (18x), *portentum* “omen”, *possessio* “seizing” (22x), *potestas* “power” (34x), *potus* “drinking” (15x), *protectio* “covering”, *prudentia* “foresight”, *redemptio* “redemption”, *regnum* “kingly power” (2x), *remissio* “returning”, *requies* “rest” (9x), *robur* “strength”, *salus* “health” (11x), *salutare* “to greet” (4x), *sanitas* “soundness of mind” (2x), *sapientia* “wisdom” (11x), *sapor* “taste”, *scientia* “knowledge” (4x), *semen* “seed” (2x), *sensus* “observation” (3x), *signum* “sign” (5x), *silentium* “silence”, *solitudo* “loneliness” (5x), *sollicitudo* “disquiet”, *somnum* “sleep” (5x), *sonitus* “sound” (2x), *spes* “hope” (4x), *splendor* “splendour”, *spiritus* “breathing”, *stupor* “stupor”, *suavitas* “sweetness”, *summitas* “summit”, *superbia* “arrogance”, *tenebrae* “darkness”, *terror* “terror” (3x), *timor* “fear” (2x), *tormentum* “torture”, *tributum*/a “payment” (2x), *tristitia* “sadness” (3x), *usura* “use”, *venia* “kindness”, *vexatio* “harassment”, *victoria* “victory”, *virtus* “virtue” (2x), *vita* “life” (8x)

Table 9.7 All SVCs in the Bible corpus for the verb *dare* “to give”.

It appears that the verb *dare* “to give” is used far more frequently than the more general *facere* “to make”.[24] The ratio is exactly 394 : 20. This is certainly related to valency, since *dare* “to give” is trivalent and the dative constituent can be used as a causee, even if a causee does not appear in all cases here. On the other hand,

24 There is an interesting parallel in the Far East language of Thai which has—apart from a lexical causative—two verbs with the same meaning as in Latin and Greek, namely one verb with the meaning of “to give” (*hây*) and another verb with the meaning “to make/to do” (*tham*). The second construction is more a permissive causative (*the seller lets the customer drink the wine = he allows the customer to drink the wine*). These verbs are combined with any other verb in a so-called serial-verb construction to express causativity (cf. Iwasaaki and Ingkaphiron 2005: 325–335).

facere "to make" has the possibility of forming a causative periphrasis with an infinitive, which, as n. 10 showed, amounts to 239. This already gives us a point for the further analysis, namely the question of whether in those cases where no causee appears, there is either a contextual ellipsis or a general constituent. This question cannot be answered here because of the vast nature of the material, but there are some other questions that remain to be answered.

5. Morphological Analysis

The first question that arises here is whether and to what extent Biblical Hebrew already possessed SVCs, apart from the verse about Sarah's laughter in (6a). Secondly, it has to be asked whether Latin simply adopts the SVC of the Septuagint or even forms some of its own.

The first question can be answered in the affirmative in view of the evidence found so far in both groups, and especially in view of the evidence in Genesis 21:6. Both translations also retain this SVC, which consists of *ʿɔsɔh* "to do, to make" and the accusative of a verbal AN. If we take the Latin SVCs from the Vulgate text as our starting point, there is a total of 106 unambiguous SVCs for *facere* "to do" and *dare* "to give" in the Old Testament where an SVC is used in Hebrew.[25] Conversely, there are 22 instances where Hebrew uses an SVC.[26]

The difference between the Vulgate and the Septuagint is evident in the different rendering of two frequently used verb forms: while Hebrew uses the Hifil verbs in most cases, namely *šɔqɔh* "to let (someone) drink" and *nuaḥ* "to let (someone) rest," Greek has two synthetic verbs, namely ποτίζω (*potízō*) "to let drink", see (49), for the first verb and παύω (*pauō*) "to finish" and compounds (καταπαύω (*katapaúō*), see (50), and ἀναπαύω (*anapaúō*)) for the second.

(49) Genesis 24:18

Ἡ δὲ εἶπεν Πίε, κύριε. Καὶ ἔσπευσεν καὶ καθεῖλεν τὴν ὑδρίαν ἐπὶ τὸν βραχίονα αὐτῆς καὶ **ἐπότισεν** αὐτόν, ἕως ἐπαύσατο πίνων.

25 Genesis 4:12; 21:6; Exodus 11:3; 12:36; Leviticus 26:3(4); 26:36; Numbers 16:14; 18:8; 18:21; 27:3(4); 27:10; 32:5; 32:29; Deuteronomy 2:5; 2:19; 4:38; 5:31; 11:14; 11:25; 21:23; Joshua 17:4a; 17:4b; 17:14; 19:49; 20:20; Judges 15:18; Ruth 4:12; 1 Samuel 2:10; 5:12; 12:17; 12:18; 14:45; 19:5; 2 Samuel 4:8; 1 Kings 5:4; 8:56; 11:11; 18:45; 21:3; 2 Kings 5:1; 8:19; 2 Chronicles 1:10; 1:12; 6:27; 17:10; 1 Esra 9:8; 9:9; 9:13; 2 Esra 4:4; 8:7; 8:11; Psalms 1:3; 2:8; 4:7(8); 13:7; 17:36; 52:7; 54:23; 60:6; 66:7; 83:12; 84:8 84:13a; 84:13b; 85:16; 110:4; 120:3; 123:6; 126:2; 131:4; 134:12; 135:21; 139:9; Proverbs 5:9; 6:4; 10:10; 13:15; Job 5:10; 38:16; 42:15; Isaiah 30:23; 42:8; 46:13; 48:11; 62:7; Jeremiah 5:24; 10:13; 14:13; 14:22; 24:9; 26:6; 34:17; 34:22; 51:16; 51:43; 51:55; Lamentations 21:18; Ez. 5:14; 29:19; 30:13; 32:23 32:25; 32:26; 47:23; Daniel 1:17; 2:21a; 2:21b; 2:37; 5:18; Zechariah 8:12; Malachi 2:2; 2:5.

26 Genesis 24:18; Numbers 5:26; Deuteronomy 3:20; 25:19; Joshua 1:13; 1 Chronicles 11:11; 22:18; 23:25; 2 Chronicles 14:7; Psalms 50:10; 79:6; 118:34; 118:73; 118:126; 118:144; 2 Samuel 7:1; 23:15; Isaiah 43:20; Jeremiah 8:14; 16:7; Ezekiel 46:18; Habakkuk 2:15.

Hē de eipen Pie, kurie. Kai espeusen kai katheilen tēn hudrian epi ton brakhiona autēs kai ***epotisen*** *auton, heōs epausato pinōn*

"And she said, Drink, lord. And she hurried and lowered the jar upon her arm and gave him a drink." (NETS, Hiebert)

(50) Deuteronomy 25:19

καὶ ἔσται ἡνίκα ἐάν **καταπαύσῃ** σε κύριος ὁ θεός σου ἀπὸ πάντων τῶν ἐχθρῶν σου τῶν κύκλῳ σου ἐν τῇ γῇ, ᾗ κύριος ὁ θεός σου δίδωσίν σοι ἐν κλήρῳ κατακληρονομῆσαι […]

kai estai hēnika ean ***katapausē*** *se kurios ho theos sou apo pantōn tōn ekhthrōn sou tōn kuklō sou en tē gē hē kurios ho theos didōsin soi en klērō kataklēronomēsai […]*

"And it shall be, when the Lord your God gives you rest from all your enemies who surround you, in the land that the Lord your God is giving you as an allotment to inherit […]." (NETS, Peters)

An SVC is possibly used in three places in the Septuagint, twice with δίδωμι (*didōmi*) "to give" and twice with ποιέω (*poieō*) "to make". In (52), the Septuagint uses δίδωμι κατάπαυσιν (*didomi katapausin*) "to let rest" instead of the synthetic verb καταπαύω (*katapauō*). In (52), the rarer AN ἔκνηψις (*eknēpsis*) "sobriety, health" is used. In (53), there is no SVC in the first clause, since here πότος (*potos*) has the meaning of συμπόσιον (*sumposion*), but ἄφεσιν ποιῶ (*aphesin poiō*) "to give rest" is clearly an SVC. These three examples clearly show that the Septuagint can also use FVG instead of a synthetic verb form.

(51) 3 Kings 8:56

Εὐλόγητος κύριος σήμερον, ὃς **ἐδωκεν κατάπαυσιν** τῷ λαῷ αὐτοῦ Ισραηλ κατὰ πάντα, ὅσα ἐλάλησεν.

Eulogētos kurios sēmeron, hos ***edōken katapausin*** *tō laō autou Israēl kata panta hosa elalēsen.*

"Blessed be the Lord today, who has given rest to his people Israel according to all that he spoke; not one word failed in all his good words which he spoke by the hand of his slave Moyses." (NETS, McLean and Taylor)

(52) Lamentations 2:18

μὴ **δῷς ἔκνηψιν** σεαυτῇ, μὴ σιωπήσαιτο, θύγατερ, ὁ ὀφθαλμός σου.

mē ***dōs eknēpsin*** *seautē, mē siōpēsaito thugater, ho ophthalmos sou.*

"Give yourself no calming down! May your eye not be silent, oh daughter!" (NETS, Gentry)

(53) Esther 2:18

καὶ ἐποίησεν ὁ βασιλεὺς πότον πᾶσι τοῖς φίλοις αὐτοῦ [...] καὶ **ἄφεσιν ἐποίησεν** τοῖς ὑπὸ τὴν βασιλείαν αὐτοῦ.

kai epoiēsen ho basileus poton pasi tois philois autou [...] kai ***aphesin epoiēsen*** *tois hupo tēn basileian autou.*

"Then the king gave a wine party for all his Friends and forces for seven days and gave rest to those under his rule." (NETS, Jobes)

In all cases, the Latin of the Vulgate has an SVC, namely *requiem dare* "to give rest" and *potum dare* "to let drink". In the three languages, we therefore see a tendency from synthetic to analytical forms. Hebrew here only has morphological forms marked by the word stem in the Hifil, Greek has morphological means, namely, for example, a verb with the suffix -ίζω *-ízō* (ποτίζω (*potizō*) "to let drink" and συνεθίζω (*sunethizō*) "to accustom someone to"), and in some cases already uses SVCs. Latin only has SVCs here because, apart from the lexemes with *-efacere* (*-ēfacere*) already mentioned and the *facere* "to make" infinitive, it has neither a morphological causative nor certain word-formation forms. So, Latin could ultimately have most SVC.

There are two syntagms that provoke an SVC in Latin, namely *bibere* "to drink" and *requiescere* "to rest" with their transitive meanings. Since it might seem easy to assume that these are the only verbs, the corpus should be checked for other verbs. There are five other verbs that cause an SVC in Latin with *dare* "to give" or *facere* "to make", but they occur only once, except only for *bijn* "to give insight, to instruct" (in the Hifil). Table 9.8 shows all 7 verbs.

bijn	Hifil: "to give insight, to instruct" (e.g. Psalms 118:25)
naḥal	Hifil: "to bequeath" (Esra 46:18)
nuwaḥ	Hifil: "to let rest" (e.g. 1 Chronicles 22:18)
šaqah	Hifil: "to let drink" (e.g. Genesis 24:18)
'ur	Hifil: "to let shine" (Esra 32:7)
yaḥal	Piel: "to give hope" (Psalms 118/119:49)
yaza'	Hifil: "to carry out" (1 Chronicles 11:14)

Table 9.8 Further Hebrew verbs with an SVC in the Vulgate.

6. Syntactic Analysis

In this part of the analysis, it is particularly important to determine the extent to which a causee appears, or, if not, whether it can be easily determined in the

context, i.e., whether there is a current ellipsis or whether it is an infinitive. In addition, the word order must be examined here, i.e., the order in which the AN and SV appear and how this can be accounted in relation to normal word order. Thirdly, it is informative to determine the type of base verbs, i.e., whether they are intransitive, transitive, or even ditransitive verbs. Apart from this word-order study, further analysis is not possible within this chapter.

Regarding the first question, it should be noted that, in most cases, a causee appears in the Latin text as in (54) and (55), which appears as a dative.

(54) Ezekiel 46:18

et non accipiet princeps de hereditate populi per violentiam [...] sed de possessione sua hereditatem dabit ***filiis suis*** *[...]*

"The prince should not take anything from the people of their inheritance [...] but he should pass on his property to his sons and daughters [...]."

(55) Psalms-G 118:49

Memor esto uerbi tui seruo tuo, in quo ***mihi*** *spem dedisti.*

"Remember your word to your servant, in which you gave me hope."

If we assume transitivity in the second question, we see that the six most common verbs include both intransitive (*requiescere* "to rest", *timēre* "to be afraid of, to terrify") and transitive verbs (*bibere* "to drink", *heredem dare* "to make someone an heir"). *In potestatem dare* "to place someone under one's authority" is also a ditransitive verb, as the base verb requires two additional complements in addition to the subject. Most verbs are intransitive and transitive base verbs.

As for the second question, of the 52 base verbs, 27 are intransitive (e.g., *bibere* "to drink", *gaudēre* "to be happy", *silēre* "to be silent"),[27] 22 are transitive (e.g., *habitāre* "to live, to dwell", *possidēre* "to own", *timēre* "to be afraid of"),[28] and three are ambi-transitive, i.e., both intransitive and transitive, namely *formidāre* "to be afraid, to terrify", *meminisse* "to remind of, to remember",

27 *Adfligere* "to afflict"; *benignum esse* "to be kind"; *bibere* "to drink"; *confidere* "to trust"; *constantem esse*/*constare* "to be constant/stable"; *dormire* "to sleep"; *fluctuare* "to fluctuate"; *frui* "to enjoy"; *gaudere* "to rejoice"; *heredem esse* "to be an heir"; *iucundum esse* "to be pleasant"; *laetari* "to rejoice"; *lucere* "to shine"; *nivem esse* "to be snow"; *oblivisci* "to forget"; *paeniteo* "to repent"; *pavescere* "to be afraid"; *pluit* "to rain"; *requiescere* "to rest"; *ridere* "to laugh"; *salvere* "to save"; *sanum esse* "to be healthy"; *sapere* "to be wise"; *silere* "to be silent"; *sobrium esse* "to be sober"; *stupere* "to be astonished"; *vivere* "to live".

28 *Benedicere* "to bless"; *capere* "to take"; *concipere* "to conceive"; *curare* "to care for"; *despicere* "to despise"; *edere* "to eat"; *habitare* "to live"; *honorare* "to honour"; *imperare* "to command"; *intellegere* "to understand"; *partiri* "to share"; *petere* "to ask for"; *placare* "to appease"; *possidere* "to own"; *protegere* "to protect"; *remittere* "to forgive"; *scire* "to know"; *sentire* "to feel"; *sperare* "to hope"; *timere* "to fear"; *tristem esse* "to be sad"; *vincere* "to conquer".

terrēre “to scare, to be afraid of”. The vast majority of verbs are both intransitive and transitive, with intransitive verbs forming a slight majority.

Finally, word order is calculated. Only the position of the AN and the SV is relevant here, as other constituents such as the subject and an adverb do not always have to appear. If, by the addition of another constituent the position is changed in a subordinate clause, such as a relative clause, this is not taken into account. A total of 360 constituents appear *in situ*, minus 17 in a subordinate clause, leaving 343 examples. These yield the following picture, Table 9.9.

AN-SV	SV-AN
57x	343
22:8%	77:2%

Table 9.9 The position of the AN and SV.

It can therefore be seen that the SV clearly appears almost three and a half times more often before the AN. This could well be motivated by semantic reasons, as the semantic core is the AN. Also, it seems to show the tendency towards V[erb]-O[bject] orders in late Latin. A similar result could be seen in the position of the causatively used *facere* “to make” + infinitive. In these instances of the Old Testament, the verb appears in 83 cases before the AN of a total of 111 instances, so in nearly 75% (Hoffmann 2014: 166).

7. Semantic Analysis

In this step of the analysis, it is first important to identify the nature of the causal relationship. It is also important to see whether there are certain semantic verb classes that are particularly prone to forming SVCs. We have already seen that *verba sentiendi*, and especially *verba timendi*, are particularly suitable for forming SVCs.

With regard to the first question, it must be clarified what kind of semantic relationship exists between the causative and the base verb. According to Haspelmath and Müller-Bardey (2004), a rough distinction must first be made between direct and indirect causatives. Within the direct type, one must determine whether it is an immediate causative; within the indirect type, whether it may be assistive or permissive—or whether these distinctions are irrelevant, as assumed by Kedar-Kopfstein (1973) in the case of the Vulgate. To this end, we start from the five most common ANs listed in Table 9.6. *Requiem dare* “to give rest”, in the literal sense of falling asleep, is a process that becomes a state in which the causer is mainly involved, but the causee less so. The same applies to *hereditatem dare* “to bequeath to”. Thus, *bibere* “to drink” and *frui* “to

rejoice" belong to the semantic class of TAKING FOOD, *requiescere* "to rest" is a verb of RESTING, *timere* "to be afraid" is a *verbum sentiendi*, while *heredem esse* "to be an heir" denotes TAKING POSSESSION. Nor are all cases controllable. This does not seem to be the case with *requiescere* "to rest" and *timere* "to be afraid". In the case of *fructum dare* "to bear fruit" and *potum dare* "to let drink", the causer is only the trigger of an action that the causee performs themselves. *Timorem dare* "to scare" is a subjective feeling that is triggered solely by the causer. In contrast, *potestatem dare* "to empower" is a process that is only initiated by the causer but continued solely by the causee. These six causative SVCs clearly show that the causative relationship can vary considerably.

As for the second question, here too we can assume the most common SVC from Table 9.6, namely *hereditas* "inheritance", *potestas* "power, permission", *fructus* "fruit", *potus* "drink", *timor* "fear", *requies* "rest", to which the following verbs belong: *heredem esse* "to be an heir", *posse*/*potentem esse* "to be able to", *frui* "to rejoice", *bibere* "to drink", *timere* "to fear", and *requiescere* "to rest". They seem to have little or nothing to do with each other.

8. Conclusions and outlook

A summary of the causatively used SVCs attested in the Bible, especially the Old Testament, shows that, even though this chapter could only examine them selectively, there are quite a few SVCs already in the Hebrew original. A prime example of this is Genesis 21:6, (28c). Compared to the 239 occurrences of *facere* "to make" with infinitive (see n. 10 above), this linguistic device is even more predominant, with about 400 occurrences in the Latin Bible.[29] Hebrew SVs are usually formed with *ʿɔsɔh* ("to do") or *nɔtan* ("to give") and the abstract noun appearing in the accusative. In Greek, the two SVs of active ποιῶ (*poiô*) "to make" and medial ποιοῦμαι (*poioûmai*) "to do" and more often δίδωμι *didōmi* ("to give") are used. In Latin, *dare* "to give" and *facere* "to make" are the preferred SVs; in the upmost cases the AN appears in the accusative, but only seldom with a preposition (in our corpus only (24), Numbers 34:5, and (36), Jeremiah 20:4).[30] Since in Hebrew there are two verb stems—mainly the Hifil verb stem (for example, *yaʾᵅkilenuʷ* "he will let us eat") and also, in the factitive subtype, the Piel verb stem—and Greek likewise has morphological means for forming special synthetic causative verbs, especially those ending in -ίζω (*-ízō*) as for instance ποτίζω (*potízō* "I let drink"), the largest number of causative

29 Though some of these instances are not used causatively.

30 Apart from three instances from the Greek and perhaps already from the Aramaic, namely (8), Matthew 6:13, and (9), Mark 14:38 parr.: both from the NT, and (10), Wisdom 15:4. This is a big contrast to German SVCs where the prepositional SVCs are prevailing. In classical Latin (see (1) and (2)), it was not so rare.

SVCs appears in the Latin of the Vulgate. In this language the two most frequent SVs are *dare* "to give" and *facere* "to make". The number of SVCs formed with *dare* "to give" is disproportionately larger (about 400 : 250), although it must be taken into account that *facere* "to make" often can express causativity with the infinitive (cf. n. 10). Compared to the Hebrew original, already the Greek language shows a significant increase. Because of the morphological resources that Latin does not have—apart from the words combined with *-efacere* as for example *mansuefacere* ("to tame")—the Vulgate as a whole has the greatest amount of causative SVCs among the three languages. So late Latin reflects more grammaticalisation than the other two languages. In using these syntagms, the language of the Vulgate did not create any new linguistic possibilities, but used a linguistic tool to express causativity, which already the Early Latin of Plautus (Rosén 1981) and the classical Latin of Cicero, Livy, and other authors (Hoffmann 1996; Pinkster 2015: 74–77) possessed to a sufficient degree.

It would be interesting to investigate the total extent of SVCs in the Vulgate in order to determine the quantitative extent of SVCs in the Vulgate compared to the number in, for example, Cicero's philosophical writings. On the one hand, one would expect a greater quantity in a translation text from a language with a morphological causative, but on the other, there is the other candidate of *facere* "to make" and infinitive, which was in classical Latin only rarely used, but arose in Late antiquity (Chamberlain 1986). As for this construction, it would be interesting to see whether there is a clear distinction between both Latin causative constructions. A parameter could be the obligatory degree of the expression of a causee which in a three-place verb like *dare* "to give" can better be expressed than in a two place verb like *facere* "to make".[31] It would also be interesting to examine the extent to which SVCs form a system in the Vulgate, or the extent to which Jerome (fourth and fifth centuries CE) continued to use this means of expression, who himself translated the Pentateuch and the psalms in the Old Testament, or at least corrected a translation close to the Septuagint. A further parameter could be the type of action (ingressive or durative) and the question, how far the Vulgate uses such distinctions which were a plus in comparison with the infinitival construction. Research into other late Latin authors would also be interesting, for instance Gregory the Great (sixth and seventh centuries CE) or Augustine of Hippo (fourth to fifth centuries CE) where one could limit oneself to a work such as *De Civitate Dei*.[32] However, such questions go beyond the scope of this chapter.

31 From a typological standpoint, it is not relevant that isolating languages like Thai use a verb which denotes GIVE to express a causative construction.

32 Cf. *qui solus … diligens et dilectus beatos facit* "[sc. God] who alone […] loves and is loved makes (people) happy" (Augustine, *City of God*, 10.17). As for Gregory the Great, the dissertation by Seitz (1938) could be useful, one of the few studies of SVCs in late Latin.

Abbreviations

AN = abstract noun

BSV = Beriger, Andreas, Widu-Wolfgang Ehlers, and Michael Fieger. 2018. *Biblia Sacra Vulgata: lateinisch—deutsch. 6 vols*. Berlin and Boston, MA: De Gruyter.

BH = *Biblia Hebraica Stuttgartensia*. Stuttgart: Deutsche Bibelgesellschaft.

Duden = Wöllstein, Angelika. 2016. *Duden: Die Grammatik*. Berlin: Dudenverlag.

LXX = Septuagint

NETS = *A New English Translation of the Septuagint* (Genesis: Robert J.V. Hiebert; Exodus: Larry J. Perkins; Numbers: Peter W. Flint; Deuteronomy: Melvin K.H. Peters; 1 (= 3) Kings: Paul D. McLean and Bernard A. Taylor; Esther: Karen H. Jobes; Psalms: Albert Piertersma; Wisdom: Michael L. Knibb; Sirach: Benjamin G. Wright; Lamentations: Peter J. Gentry;). Electronic Edition 2009. Second printing (https://ccat.sas.upenn.edu/nets/edition/).

NIV = New International Version [of the English Bible]. Online Version (https://www.biblegateway.com/versions/New-English-Translation-NET-Bible/).

SV = support verb

SVC = support-verb construction

Bibliography

Ballweg, Joachim et al. (eds.). 1997. *Grammatik der deutschen Sprache*. Vol. 2. Berlin: De Gruyter.

Bekins, Peter. 2014. *Transitivity and Object Marking in Biblical Hebrew: An Investigation of the Object Preposition 'et*. Winona Lake, IN: Eisenbrauns. https://doi.org/10.1163/9789004370166

Biville, Frédérique. 1995. Enoncés factitifs en latin. Syntaxe et sémantique. In Dominique Longrée (ed.), *De Usu*, 31–44. Louvain-la-Neuve: Peeters.

Bortolussi, Bernard, and Lecaudé Peggy (eds.). 2014. *La Causativité en latin*. Paris: L'Harmattan.

Chamberlain, Jeffrey T. 1986. *Latin Antecedents of French Causative Faire*. Frankfurt/Main: Peter Lang.

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. Oxford: Blackwell.

Dixon, Robert 2000. A Typology of Causatives. Form, Syntax and Meaning. In Robert Dixon and Alexandra Aikhenvald. *Changing Valency: Case Studies in Transitivity*, 30–83. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511627750.003

Edwards, Henry 1917. *Caesar: The Gallic War.* Cambridge, MA: Harvard University Press.

Fischer, Bonifatius. 1977. *Novae Concordantiae Bibliorum sacrorum iuxta vulgatam versionem critice editam quas digessit Bonifatius Fischer OSB*. Stuttgart-Bad Cannstatt: Frommann & Holzboog.

Fischer, Bonifatius, Robert Weber, and Roger Gryson. 1994. *Biblia sacra: iuxta vulgatam versionem*. Stuttgart: Deutsche Bibelgesellschaft.

Grimm, Wilhelm, and Jacob Grimm. 2004. *Deutsches Wörterbuch: Der digitale Grimm. Bearb. v. H.W. Bartz*. Frankfurt a.M.: Zweitausenundeins.

Happ, Heinz. 1976. *Grundfragen einer Dependenzgrammatik des Lateinischen*. Göttingen: Vandenhoeck and Ruprecht.

Haspelmath, Martin, and Thomas Müller-Bardey. 2004. Valency Change. In Geert Booij et el. (eds.), *Morphologie—Morphology*, *vol. 2*, 1130–1145. Berlin and New York: De Gruyter. https://doi.org/10.1515/9783110172782.2.14.1130

Helbig, Gerhard, and Buscha Joachim. 1989. *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht*. Leipzig: Volkseigener Betrieb Verlag Enzyklopädie.

Hetzron, Robert. 1990. Hebrew. In Bernard Comrie (ed.), *The World's Major Languages*, 686–704. Oxford: Oxford University Press.

Hoffmann, Roland. 1996. Funktionsverbgefüge im Lateinischen. In Alfred Bammesberger and Friedrich Heberlein (eds.), *Proceedings of the Eighth International Colloquium on Latin Linguistics*, 200–212. Heidelberg: Winter.

— 2005. Functional Verb Constructions of the Type *adferre* + accusative. Synchronic and Diachronic Observations. In Gualtiero Calboli (ed.), *Papers on Grammar Ix.2: Proceedings of the Twelfth International Colloquium on Latin Linguistics (Bologna, 9-14 June 2003), vol. 2*, 547–559. Rome: Herder Editrice.

— 2008. Causative Constructions in Late Latin Biblical Translations. In Roger Wright (ed.), *Latin vulgaire—latin tardif VIII. Actes du VIIIe colloque international sur le latin vulgaire et tardif, Oxford, 6–9 septembre 2006*, 160–172. Hildesheim: Olms-Weidmann.

— 2014. Les constructions causatives dans les traductions latines des textes hébreux et grecs : le cas de la Vulgate de Jérôme. In Bernard Bortolussi and Lecaudé Peggy (eds.), *La Causativité en latin*, 143–176. Paris: L'Harmattan.

— 2018. Criteria for Describing Valency in Latin Function Verb Constructions. In Colette Bodelot and Olga Spevak (eds.), *Les constructions à verbe support en latin*, 75–93. Clermont-Ferrand: Presses Universitaire Blaise Pascal.

— 2023. Latin Support Verb Constructions. A View from Language Typology. In José Miguel Baños Baños, María Dolores Jiménez López, María Jiménez Martínez, and Cristina Tur (eds.), *Collocations in Theoretical and Applied Linguistics: From Classical to Romance Languages*, 21–56. Madrid: Guillermo Escolar.

Iwasaaki, Shoichi, and Preeya Ingkaphiron. 2005. *A Reference Grammar of Thai*. Cambridge: Cambridge University Press.

Jenni, Ernst. 1992. Kausativ und Funktionsverbgefüge, Sprachliche Bemerkungen zur Bitte: „Führe uns nicht in Versuchung". *Theologische Zeitschrift* 48. 77–88.

Jeremias, Joachim. 1971. *Neutestamentliche Theologie*. Göttingen: Vandenhoeck & Ruprecht.

Kedar-Kopfstein, Benjamin. 1973. Die Wiedergabe des hebräischen Kausativs in der Vulgata. *Zeitschrift für die alttestamentliche Wissenschaft* 85. 196–219.

Keyes, Clinton 1928. *Cicero: On the Republic, On the Laws.* Cambridge, MA: Harvard University Press.

Luz, Ulrich. 2002. *Das Evangelium nach Matthäus. Bd.1: Mt. 1-7*. Zürich: Benziger.

Madrigal Acero, Lucía. 2024. What Can Be Used in Greek and Latin? A Comparative Study of the Support Verbs χράομαι khraomai and utor. In Victoria B. Fendel (ed.), *Support-verb Constructions in the Corpora of Greek: Between Lexicon and Grammar*, 65–91. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.14017925

Mayer, Rudolf. 1965. *Einleitung in das Alte Testament, 1. Teil: Allgemeine Einleitung*. Munich: Max Hueber.

Pinkster, Harm. 2015. *Oxford Latin Syntax. Vol.1: The Simple Clause*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199283613.001.0001

Rahlfs, Alfred. 1935. *Septuginta id est Vetus Testamentum Graecae iuxta LXX interpretes.* Stuttgart: Priviligierte Württembergische Bibelanstalt.

Rosén, Hannah. 1981. *Studies in the Syntax of the Verbal Noun*. Munich: Fink-Verlag.

Schwarz, Günther. 1987. "Und Jesus sprach": Untersuchungen zur aramäischen Urgestalt der Worte Jesu. Stuttgart: Kohlhammer.

Seitz, Johannes. 1938. *Über die Verwendung der Abstrakta in den Dialogen Gregors des Großen.* Borna-Leipzig: Noske.

Song, Jae Jung. 2005a. Periphrastic Causative Constructions. In Martin Haspelmath et al. (eds.), *World Atlas of Language Structures*, 446–449 (Ch. 110). Oxford: Oxford University Press.

— 2005b. Nonperiphrastic Causative Constructions. In Martin Haspelmath et al. (eds), *World Atlas of Language Structures*, 450–453 (Ch. 111). Oxford: Oxford University Press.

Steurer, Rita Maria. 1989-2003. *Das Alte Testament. Interlinearübersetzung Hebräisch-Deutsch nach der Biblia Hebraica Stuttgartensia. Vol.1-Vol.4*. Neuhausen and Stuttgart: Hänssler.

Van Pottelberge, Jeroen. 2007. Funktionsverbgefüge und verwandte Erscheinungen. In Harald Burger et al. (eds.), *Phraseology. An International Handbook of Contemporary Research, vol. 1*, 436–444. Berlin: De Gruyter. https://doi.org/10.1515/9783110171013.436

10. Understanding the expression of pain across time: Metaphor and intensity in Latin (and Spanish) support-verb constructions with *dolor* (2nd c. BCE to 15th c. CE)

María Isabel Jiménez Martínez

1. Introduction

In this study, we analyse support-verb constructions (SVCs henceforth) with the noun *dolor* from a diachronic perspective, using a corpus that spans from archaic and classical Latin (second century BCE to the second century CE), through late Latin in southern Europe (seventh to the ninth century CE)—to medieval Spanish (thirteenth to the fifteenth century CE).[1] Our aim is to identify the constructions and metaphorical patterns that encode the experience of pain, particularly its intensity, and to examine how these patterns are preserved, reshaped, or innovated over time. To this end, the research addresses the following questions:

1. How is the experience of pain linguistically structured in classical Latin, and how do these patterns evolve into late Latin and medieval Spanish?
2. How is pain conceptualised at the semantic and syntactic levels through SVCs?
3. Which metaphorical networks are activated in these constructions, and how do they contribute to the encoding of intensity?

1 The dataset is openly accessible here: https://doi.org/10.5281/zenodo.19055091

©2026 María Isabel Jiménez Martínez, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.10

The chapter is structured as follows: Section 2 reviews the state of the art, discussing the concept of SVCs, particularly their intensive extensions, and their relationship to metaphor theory. Section 3 presents the methodology and describes the corpora, including how examples were selected and coded. In Section 4, we examine the SVCs identified in the classical Latin corpus, analysing their semantic types and underlying metaphors. Section 5 provides a parallel analysis of SVCs in late Latin and medieval Spanish, comparing them with those in Latin in terms of diversity, complexity, and metaphorical strategies. Finally, we summarise the main findings, highlighting the cognitive and cultural implications of how pain is conceptualised in these languages over time in Section 6.

2. State of the art

2.1 Support-verb constructions: A definition

In this study, we understand SVCs as multi-word expressions formed by a noun and a verb that jointly occupy the predicate slot of a sentence; accordingly, they are understood at the lexicon-syntax interface. In SVCs, the noun—always predicative—serves as the semantic and syntactic nucleus of the construction, providing its argument structure,[2] while the support verb acts as a structural frame, supplying person, tense, number, voice, and mood (Baños and Jiménez López 2024: 94ff.). Given the semantic primacy of the noun, it also constrains the range of verbs with which it may combine.

Under this definition, the most prototypical SVCs are multi-word combinations in which the verb has a general or weak lexical meaning, as *have, make,* or *give*, and a subject that is coreferential with the first argument of the predicative noun, for example, *she made a decision*, where *she* is both the syntactic subject of *made* and the semantic agent of *decision*. However, in these combinations we also include constructions in which the verb carries fuller lexical content (Baños and Jiménez López 2024: 94ff.). In these cases, the verb is metaphoric (and constitutes a lexical alternative to a general support verb, as illustrated by English *shoot an answer* which functions as a lexical substitute for *give an answer*) or/and contribute aspectual (inchoative or terminative), diathetic (causative or passive), or intensive nuances. The syntactic configuration inside SVCs may vary: the noun may appear as the object, as in English *make a decision*, as the subject, as in Latin, *cupiditas nascitur* "desire arises", or even as a locative complement, as in Spanish *entrar en razón* "to come to one's senses".[3]

2 Here we adopt the terminology of Functional Grammar adapted to Latin (Baños 2023).
3 The definition of SVC adopted here is broader than that commonly found in the literature

Among SVCs, aspectual and diathetic extensions have been widely studied in Latin and the Romance languages (Gross 2022; Vivès 1983; Cicalese 1999; Mendózar Cruz 2020; Salas 2024). By contrast, intensive SVCs, those that highlight the degree or force with which an event unfolds, have received far less attention (Gavriilidou 2004), and these are of particular interest here.

As Gavriilidou (2004: 298) notes, languages have multiple strategies to encode intensity: morphological, for example, intensive prefixation, suffixation, or compounding, prosodic, for example, stress, emphasis, and intonation, and syntactic, including intensive adjectives, such as Latin *magnum*, intensive adverbs, such as Latin *ualde*, and intensive SVCs.

In this study, we distinguish two mechanisms through which SVCs code intensity: (i) *degree constructions*, which encode an increase or decrease in the intensity of the event, (1a), and (ii) *intensive extensions,* which occur in SVCs of any semantic type (general, aspectual, or diathetic) and add an explicit nuance of intensity, as illustrated in (1b).

(1a) Cicero, *Ad familiares,* 4.5.6

nullus ***dolor*** *est quem non longinquitas temporis* ***minuat*** *ac* ***molliat***.

“There is no grief that is not diminished and mitigated by the lapse of years.”

(1b) J. Fernández de Heredia, *Historia Troyana* (1376–1396 CE)

fue turbado *grandemente et* ***atormentado de grant dolor.***

“He was greatly disturbed and tormented by intense pain.”

In sum, the concept of SVC used in this study includes not only support verbs in the strict sense but also their extensions. This broader approach is necessary because it captures the full combinational behaviour of the predicative noun *dolor* and helps explain why they combine with particular verbs.

2.2 Support-verb constructions and metaphor

The almost automatic association we establish between concrete and abstract realities through metaphor constitutes one of the cognitive foundations of

(see, e.g., Bodelot and Spevak 2018), where such constructions are restricted to combinations involving prototypical support verbs, i.e. verbs that do not contribute independent semantic content but merely function as a syntactic support for the predicative noun. From our perspective—also adopted in Baños et al. (2022)—these constructions include not only proper support verbs but also other verbs, typically of a metaphorical nature, which can be analysed as extensions of support verbs. In such cases, the verb may serve as a lexical alternative to the canonical support verb, e.g. Spanish *dar una bofetada* “give a slap” vs. *soltar una bofetada* “let loose a slap”, or it may contribute aspectual or diathetic information to a more generic support-verb construction.

human thought, enabling us to organise, interpret, and communicate our experiences. This perspective, developed in the 1980s by scholars such as Lakoff, Johnson, and Langacker (Lakoff and Johnson 1980; Lakoff 1987, 1993; Langacker 1987, 1991; Kövecses 2000, 2002) understands metaphor as a mechanism that systematically maps a concrete, literal source domain onto a more abstract target domain (as in TIME is MONEY).

SVCs provide a particularly insightful site for exploring the interaction between grammar and metaphor. As noted, an SVC combines an abstract noun with a verb to form a single predicate, in which the noun carries the core meaning and the verb adapts its sense accordingly, often in a metaphorical way. For instance, in Livy (first century BCE/first century CE), the SVC *concoquere consilia* does not refer to literal cooking but to careful deliberation, because in Latin—as in many other languages—the domains of cooking and digestion function as metaphorical frameworks for thinking, conceptualised as a slow, elaborate, and transformative process. Crucially, this noun-verb combination reflects a conventionalised metaphorical pattern, systematically realised within the SVC.

Metaphorical SVCs[4] are, therefore, multi-word combinations in which the meaning arises from the interaction of a metaphorical verb with the noun. This noun-verb pairing, which exhibits semantic interdependence and motivation, evokes a well-established metaphorical pattern within a given linguistic community, often recognised across languages. In other words, these are concrete lexical combinations that activate and reinforce familiar metaphorical patterns, allowing speakers to convey and understand complex concepts through more concrete experiences, as in THINKING is COOKING.

Metaphorical SVCs vary in their degree of conventionalisation and frequency. Some, like English *keep a secret*, are highly lexicalised and frequent, while others appear rarely in speech and corpora, like Latin *dolor tenebrat* "pain darkens". Yet both qualify as SVCs, as they meet the definitional criteria and function as potentially metaphorical expressions within the same underlying pattern, reflecting the broader system of metaphorical cognition they encode.

3. Methodology and corpus

The analysis in this study is based on the target-oriented methodology (Stefanowitsch 2006; Deignan 2005; Buccheri et al. 2021), which examines linguistic metaphors by starting from the abstract nouns most closely associated

4 The convergence of these two linguistic phenomena gives rise to the notion of metaphorical collocation, that is, the conventional association of metaphorical meanings with specific lexical combinations. Although frequently invoked, the concept has rarely been defined with precision in the literature. Patekar (2022) and Malik et al. (2025) represent recent efforts to fill this theoretical gap.

with a target domain. In our case, the domain is "pain", in archaic and classical Latin (second century BCE to the second century CE), late Latin (seventh to ninth centuries CE) and medieval Spanish (thirteenth to fifteenth centuries CE), using *dolor* as the keyword.

This approach has clear advantages. As Fedriani (2025: 219 ff.) notes, it allows for the systematic identification and classification of metaphors in large corpora, enabling both quantitative and qualitative comparisons while avoiding reliance on researcher intuition. Its main limitation is that it only captures metaphors triggered by the chosen term, excluding other expressions of the concept that use different words, alternative metaphorical patterns, or discursive contexts without the lexeme (Fedriani 2025: 219). These forms, though outside the scope of this study, remain relevant for a full understanding of how pain is conceptualised.[5]

With this limitation in mind, we created three separate corpora:

1. The archaic and classical Latin corpus: we analysed all the contexts in which the noun *dolor* appears in its various forms in seven authors representing different periods, selected from the Latin corpus of the *Packard Humanities Institute*.[6] This process yielded a total of 640 attestations.
2. The late Latin corpus: we studied all occurrences of the term *dolor* between 650 and 850 CE across authors from the Iberian Peninsula, the Italian Peninsula, and the territory of present-day France, as provided by the *Corpus Corporum.* In total, we collected 542 examples.
3. The medieval Spanish corpus: we extracted the data from the *Corpus Diacrónico del Español* (RAE), compiling a body of 600 examples in which the term *dolor* appears, evenly distributed across the 13th, 14th, and 15th centuries, with 200 randomly selected examples from each century.

This selection provides a representative sample of the various historical and linguistic contexts in which the concept of "pain" emerges, allowing us to analyse both the metaphors and the figurative patterns associated with this emotion in both languages.[7]

5 Therefore, due to methodological limitations, this study does not aim to capture the general metaphorical conceptualisation of pain. Instead, it focusses on the subset of metaphorical expressions occurring as SVCs which include the noun *dolor* in the corpora. While this approach captures only part of the picture, it provides meaningful insight into how pain was historically conceptualised, revealing patterns in Latin and medieval Spanish and offering a window into the evolution of this concept over time.

6 We include Plautus as a representative of the archaic period; Cicero (*In Verrem* and *Epistulae ad familiares*), Ovid (*Metamorphoses*), Vergil (*Aeneid*, *Eclogues*, and *Georgics*), and Livy (*Ab urbe condita*) for classical Latin; Seneca (all his tragedies) for post-classical Latin; and Celsus (*De Medicina* I–II) for late Latin.

7 However, we are aware that the composition of the corpora is not entirely comparable, as they consist of texts from very different genres, themes, dimensions, and sociolinguistic contexts.

Each sentence was manually analysed by one annotator,[8] who identified all the SVCs present across the three corpora. Table 10.1 shows the distribution of these constructions by author, totalling 166 SVCs in archaic and classical Latin, 201 in late Latin and 293 in medieval Spanish. Finally, we classified these constructions into four groups: general, aspectual, causative, and degree SVCs.

SVC	Archaic- classical Latin 2nd c. BCE to 2nd c. CE	Late Latin 650–850 CE	Medieval Spanish 850–1450 CE
General	36 (21.7%)	85 (42.3%)	197 (67.2%)
Aspectual	70 (42.2%)	72 (35.8%)	31 (10.6%)
Causative	43 (25.9%)	34 (16.9%)	42 (14.3%)
Degree	17 (10.2%)	10 (5%)	23 (7.9%)
Total	166	201	293

Table 10.1 Distribution of pain SVCs in our corpora.

Finally, following the line of work developed by Buccheri et al. (2021), Fedriani and De Felice (2022), and Leotta (2023), we identified the metaphorical schemas present in the SVCs. To do so, we adopted the typical formulation used in cognitive linguistics, of the type "X is Y," as in PAIN is a WILD ANIMAL or PAIN is FIRE. For instance, in (2) Eugenius Toletanus (seventh century CE) represents pain (the target domain) as an enemy (the source domain):

(2) Eugenius Toletanus, *Praefatio* (657 CE)

Hinc ***dolor*** *ossa* ***terit****, cor pauor inde quatit*.

"From here, pain gnaws at the bones; from there, fear shakes the heart."

Consequently, any conclusions drawn must be approached with caution.

8 Not all examples are clearly identifiable as SVCs, nor do they always reveal easily their metaphorical patterns. To annotate the corpus, we took the following steps. We considered constructions to be SVCs only if: (i) they consist of a verb and a noun dolor [morphology]; (ii) the noun occupies the syntactic position of the subject, direct object, or the prepositional complement of the verb [syntax]; and (iii) the noun functions as the semantic head of the construction, while the verb expresses (a) the mere experience of pain (general SVCs), (b) its onset or cessation (aspectual), (c) the causes triggering or relieving it (causative), or (d) the increase or decrease of the sensation (degree). Therefore, examples such as *dolorem miscere* "to mix pain" or *mostrar el dolor* "to show pain" are excluded from the corpus. Regarding metaphor annotation, examples are assigned to a metaphorical pattern only when the noun-verb combination evoked a clear image or the verb was etymologically linked to it.

It is important to consider the argument structure of the noun under study. Like all psychological predicates, *dolor* involves two semantic actants: an entity that undergoes the feeling (the experiencer) and a stimulus or source of that feeling (the cause). As we will see, in the various SVCs these actants take on different syntactic realisations.

4. *Pain* in archaic and classical Latin

Based on the analysis of the data for the archaic and classical Latin noun *dolor*, we have been able to outline the network of SVCs in which it appears, most of them metaphorical in nature and many carrying an intensifying function, as illustrated in the following examples (3):

(3a) Celsus, *De Medicina*, 1.3.20

isque, quibus, si per plures dies non uomerunt, ***dolor*** *praecordia* ***infestat****.*

"as well as in the case of those who become troubled by pain over the heart when they have not vomited for several days."

(3b) Cicero, *Ad familiares*, 5.17.4

…sin autem omnia quae sentio perscripserim, ne ***refricem*** *meis litteris desiderium ac* ***dolorem tuum****.*

"…whereas if I write down all I feel, my letter may exacerbate the pain of your separation."

The SVC *dolor infestat* "pain infests (sb.)" in example (3a) is inchoative, as it focusses on the initial moment in which the sensation begins to be felt. To convey this aspectual meaning, Celsus draws on the metaphor PAIN is an OPPONENT. Here, pain is conceptualised as an adversary launching an attack on the body (*praecordia* "diaphragm"). The intensity of the sensation is, thus, expressed through a familiar and sensorially recognizable schema—war—which highlights both its force and its aggressiveness.

Since pain is a gradable psychological state, its collocational range includes verbs that express both increases and decreases in intensity. This is precisely the case of the SVC *dolorem refricere* "to re-rub pain" in example (3b), which instantiates the metaphor PAIN is an OBJECT. In this construction, the intensification of pain is represented through the manipulation of the metaphorical object: to "re-rub" the pain implies heightening it, so the idea of "rubbing it strongly" becomes a linguistic image that conveys an especially forceful and intense experience of the emotion.

As we can see, these constructions are inherently intensifying, which is

why we regard them as extensions of SVC. A construction can be considered an intensifying extension of an SVC when, in addition to a neutral form, there exists an alternative that expresses the same basic meaning but does so with a stronger or more forceful degree of intensity (Tur 2019: 178 ff.).

In what follows, we analyse the interconnections among the SVCs in which the noun *dolor* appears, the metaphors that operate within them, and the different degrees of intensity with which this sensation is expressed.

4.1 Support-verb constructions with *dolor* "pain"

Table 10.2 presents the set of SVC in which the noun *dolor* appears in our corpus, organised into four groups according to their semantic value: general SVCs, aspectual SVCs (inchoative and terminative), causative SVCs (also inchoative and terminative), and degree SVCs (expressing increase or decrease). For each category, we include in Table 10.2 the main constructions attested in the texts at least twice, although we have also counted in the total those that appear only once. Finally, each SVC is listed together with the syntactic formalisation of its actants: the experiencer (Exp.) and the cause (Caus.).[9]

Types of SVC		SVC	Tokens	Total
General		*aliquis* Exp. *dolorem fert* "sb. brings sb. pain"	8	36
		aliquis Exp. *dolorem habet* "sb. has/is in pain"	4	
		aliquis Exp. *dolore est* "sb. is in pain"	3	
		dolor exardet "pain burns"	2	
		Others	19	
Aspectual	Inchoative	*dolor [ex/co/ob]oritur* "pain arises"	8	55
		aliquis Exp. *dolorem capit* "sb. receives pain"	4	
		dolor [e]vincit "pain subdues"	4	
		dolor commovit "pain moves"	3	
		dolor tangit "pain touches"	2	
		dolor accendit "pain kindles"	2	
		Others	32	
	Terminative	*aliquis* Exp. *dolorem finit* "sb. ends pain"	2	15
		Others	13	

9 Table 10.2 lists the SVCs that occur at least twice in our corpus. The category "Others" groups together the SVCs that appear only once, in order to improve the readability of the table.

Types of SVC		SVC	Tokens	Total
Causative	Inchoative	*aliquis* Caus. *aliquem* Exp. *dolore adficit* “sb. affects sb. with pain”	10	28
		aliquis Caus. *alicui* Exp. *dolorem affert* “sb. brings pain to sb.”	3	
		aliquis Caus. *dolorem movet* “sb. stirs pain”	3	
		aliquis Caus. *alicui* Exp. *dolorem inurit* “sb. burns sb. with pain”	2	
		Others	10	
	Terminative	*aliquis* Caus. *dolorem levat* “sb. lifts pain”	8	15
		aliquis Caus. *dolorem pellit* “sb. strikes pain”	2	
		Others	5	
Degree	Plus	*dolor [in]crescit* “pain grows”	6	12
		dolor furit “pain rages”	2	
		aliquis Caus. *alicui* Exp. *dolorem auget* “sb. increases pain for sb.”	2	
		Others	2	
	Minus	*aliquis* Caus. *alicui* Exp. *dolorem minuit* “sb. reduces pain for sb.”	2	5
		aliquis Exp. *dolorem lenit* “sb. soothes pain”	2	
		Others	1	
Total		104 SVC types		166

Table 10.2 Distribution of SVCs with *dolor* “pain” in archaic and classical Latin.

From the information presented in Table 10.2, several conclusions can be drawn that offer an overall view of the collocational behaviour of this noun. From a quantitative perspective, *dolor* proves to be a quite collocational noun: one quarter of all its occurrences in the authors examined (166 out of the 640 contexts analysed) appear in SVCs. These 166 examples are distributed across 104 SVCs, which highlights the remarkable versatility and creativity surrounding this concept.

The highest productivity, in both diversity and frequency, is found in the inchoative aspectual constructions, which account for 55 examples across 38 SVCs. These are followed by the general constructions expressing the basic idea of “experiencing/suffering pain” (36 examples in 23 SVCs), and by the inchoative causatives (28 examples in 14 SVCs). This distribution shows that the culturally and cognitively most salient elements in the conceptualisation of pain are its

onset—attested in 83 examples—and, secondly, the experience of suffering itself. Latin, therefore, structures the narrative of pain by emphasising its emergence. By contrast, combinations referring to the end of pain are scarce: only 30 in total (15 terminative aspectual constructions and 15 terminative causatives).

Five of these SVCs appear more than five times in our corpus and belong to different semantic types: the general SVC *dolorem ferre* "to suffer pain" (4a); the inchoative construction *dolor [ex-/ob-/co-]oritur* "pain arises" (4b); the causative constructions *dolore adficere* "to inflict pain" (inchoative) (4c) and *dolorem levare* "to alleviate pain" (terminative) (4d); and finally, the degree SVC *dolor [in]crescit* "pain grows" (4e).

(4a) Ovid, *Metamorphoses*, 11.328-329

quam miser amplexans ego tum patriumque ***dolorem*** */ corde* ***tuli*** *fratrique pio solacia dixi.*

"Wretched, I embrace her, feeling her father's grief in my heart, and to my dear brother I spoke words of comfort."

(4b) Plautus, *Amphitruo*, 1091-1093

Postquam parturire hodie uxor occepit tua, / ubi utero ***exorti dolores****, ut solent puerperae / inuocat deos inmortales.*

"After she began to feel near her time to-day and her pains were setting in, she called on the immortal gods to help her."

(4c) Cicero, *Ad familiares*, 10.7.2

Non me dius fidius mediocri ***dolore adficiebar*** *cum alii occupare possessionem laudis uiderentur*.

"I have felt no small chagrin, I give you my word, when others appeared to be staking out claims to glory ahead of me."

(4d) Cicero, *Ad familiares*, 5.16.1

…sed adhibere aliquam modicam consolationem, ***quae leuare dolorem tuum*** *posset si minus sanare potuisset*.

"…and require me to proffer a measure of comfort, which may alleviate, if it cannot heal, your pain."

(4e) Senenca, *Medea*, 951

rursus ***increscit dolor*** */ et feruet odium…*

"Once more my pain grows and my hatred burns."

4.1.1 General support-verb constructions

Indeed, within the group of general constructions—or SVCs in the strict sense—characterised by the use of a verb with a broad meaning and by their semantic equivalence to the simple verb *doleo* "to feel pain, to suffer", the construction *dolorem ferre* "to suffer pain", which conveys the idea that the experiencer endures or bears the pain, is the option most favoured by the authors in the corpus, (4a). Alongside this construction, the next most frequent general SVCs are also formed with verbs of a broad and polysemous meaning, such as *habere* "to have", (5a), and *esse* "to be", (5b):

(5a) Celsus, *De medicina*, 2.7.8.1

Si mulier a partu vehementes ***dolores habet***.

"If after childbirth a woman has severe pains."

(5b) Cicero, *Ad familiares*, 6.4.4

Quanto ***fuerim dolore*** *meministi.*

"You remember how sorely distressed I was."

In addition to these, the analysed contexts contain other SVCs that likewise predicate the experience of pain but introduce an element of intensity. For this reason, we classify them as intensifying extensions of the general SVCs. Among the most notable are *dolor deuorat* "pain consumes", (6a), *dolor feruet* "pain blazes", (6b), and *dolor incubat* "pain reclines", (6c):

(6a) Ovid, *Metamorphoses*, 13.538–540

Troades exclamant, obmutuit illa dolore, / et pariter vocem lacrimasque introrsus obortas / ***deuorat ipse dolor***.

"The Trojan women shrieked at the sight; but she was dumb with grief; her very grief engulfed her powers of speech, her rising tears."

(6b) Seneca, *Phoenissae*, 352

timet animus ira, ***feruet immensum dolor***.

"My spirit is swelling with anger, my pain is burning beyond measure."

(6c) Seneca, *Phaedra*, 99

Sed maior alius ***incubat*** *maestae* ***dolor***.

"But another, greater pain weighs on my distress."

As can be observed, the syntactic expression of these constructions is also variable. The most common pattern is for the experiencer to appear as the grammatical subject, as in (4a) and (5); however, the way "pain" is encoded differs across constructions. In *dolorem ferre* "to suffer pain", (4a), and *dolorem habere* "to have pain", (5a), the noun functions as the object of the verb, whereas in *dolore esse* "to be in pain", (5b), it appears in the ablative case as a complement of manner (Mendózar Cruz 2019: 190). In all the constructions in (6), finally, "pain" is expressed as the subject of the construction in the nominative case.

4.1.2 Aspectual support-verb constructions

Among the aspectual SVCs involving *dolor*, those that focus on the onset of the sensation are the most frequent. In addition to *dolor [ex-/co-/ob-]oritur* "pain arises", (4b), highly productive constructions are formed with haptic verbs, such as *dolorem capere* "to take pain", (7a), or *dolorem accipere* "to receive pain", (7b) in which the experiencer is conceptualised as the grammatical subject and "pain" as the object that is *taken*. These combinations can be understood as the inchoative counterpart of the general SVCs constructed with possession verbs: the relation *dolorem capere—dolorem habere* "to take pain—to have pain" reflects the conceptual transition from the acquisition of pain to its possession by the experiencer.

(7a) Plautus, *Truculentus*, 455–456

quanta est cura in animo, quantum corde ***capio*** */* ***dolorem****.*

"What anxiety of mind, what agony of heart I feel."

(7b) Cicero, *In Verrem*, 2.4.135

verum illud est quam ob rem haec commemorem, quod existimare uos hoc uolo, mirum quondam ***dolorem accipere*** *eos ex quorum urbibus haec auferantur.*

"my purpose in mentioning these few is to convince you that an extraordinary degree of pain has been caused to those whose towns have been robbed of such treasures."

Another group of verbs that frequently occurs in inchoative constructions is that of motion. On the one hand, transitive verbs such as *commouere* "to move" participate, giving rise to constructions in which the subject is the noun *dolor* and the displaced object is the person or being who experiences it, (8):

(8) Cicero, *Ad familiares*, 4.5.2

Quid est quod tanto opere ***te commoueat tuus dolor intestinus****?*

"What reason is there why your domestic sorrow should affect you so sorely?"

On the other hand, in these SVCs we also find intransitive telic motion verbs. In these cases, as we will discuss later (Section 4.2.2), the idea of reaching the endpoint of a movement is used to express the onset of a state. This occurs in *dolor tendit* "pain extends", (9a), and *dolor transit* "pain passes" (9b). In these instances, the grammatical subject of the SVC is the pain itself:

(9a) Celsus, *De medicina*, 2.8.39

Difficilis aeque curatio est in eadem aetate, ubi umerum ***dolor uel ad manus pervenit uel ad scapulas tendit***.

"Treatment is likewise difficult at that age when pain in the shoulders either spreads to the hands or extending to the shoulder-blades."

(9b) Celsus, *De medicina*, 2.7.22

Si a coxis et inferioribus partibus ***dolor in pectus transit*** *[…] suppurationis eo loco periculum est*.

"There is danger of suppuration in the chest when pain spreads there from the hips and lower parts."

Regarding terminative aspectual SVCs, there is a notable formal diversity, although their frequency is very low: all of them are attested only once, except for *dolorem finire* "to finish pain", (10a), which appears twice. Within this group, some constructions—such as the one just mentioned—have the experiencer as the subject and "pain" as the object; in others, by contrast, *dolor* occupies the position of the grammatical subject, as in *dolor fugit* "pain flees", (10b), referring to an entity with its own agency, capable of deciding to withdraw from the pain sufferer:

(10a) Vergil, *Aeneid*, 12.880-881

possem ***tantos finire dolores*** */ nunc certe, et misero fratri comes ire per umbras!*

"Then I could end this anguish, and pass through the shadows at my poor brother's side!"

(10b) Vergil, *Aeneid*, 12.422

subitoque omnis de corpore ***fugit*** *quippe* ***dolor***.

"and suddenly, in truth, all pain fled from the body."

4.1.3 Causative support-verb constructions

The most productive SVC of our corpus is *dolore adficere* "to affect sb. with pain" with a total of ten examples. This is a causative construction with inchoative

value, as it can be paraphrased as "to make someone begin to feel pain." In this, the cause of the pain is expressed in the nominative case, the experiencer in the accusative, and "pain" in the ablative, as it is understood as the instrument that inflicts the pain. Example (4c) above and (11) below clearly illustrate this syntactic configuration:

(11) Cicero, *In Verrem*, 2.5.123

an ***uos*** *quoque* ***hic acerbissimus innocentium cruciatus et maeror*** *pari* ***sensu doloris adficit****?*

"Or are the agonies of pain and grief inflicted on these innocent persons causing yourselves equal distress?"

Other causative SVCs with this value among the authors studied are *dolorem afferre* "to bring pain" (12a) and *dolorem mouere* "to stir up pain" (12b). In these constructions, the subject is an entity that causes the pain, the "pain" appears as the object of the verb, and the experiencer is expressed in the dative case:

(12a) Cicero, *Ad familiares*, 16.21.2

tantum enim mihi ***dolorem*** *cruciatumque* ***attulerunt*** *errata aetatis meae.*

"Young men make mistakes, and mine have brought me so much unhappiness and torment."

(12b) Ovid, *Metamorphoses*, 5.401

haec quoque uirgineum ***mouit*** *iactura* ***dolorem****.*

"the loss of her flowers even at such a time aroused new grief."

We have also identified causative SVCs with terminative value, that is, constructions in which an entity acts as the cause and brings about the cessation of the pain experienced by the experiencer. This group includes *dolorem leuare* "to alleviate pain", (4d), as well as other constructions that follow the same semantic-syntactic pattern, such as *dolorem sanare* "to cure pain", (4d), and *dolorem comprimere* "to compress pain".

4.1.4 Degree support-verb constructions

Finally, the noun *dolor* participates in SVCs that predicate the increase or decrease of the sensation, which are organised into three distinct syntactic patterns. First, there are constructions in which *dolor* functions as the grammatical subject and the verb expresses the intensification of the sensation, as in *dolor [in]crescit* "pain grows", (4e); in these cases, the emotion is identified with a living being, that is developing its body. Second, there are SVCs in which the experiencer occupies

the subject position and "pain" is encoded as the object; this is the case with *dolorem lenire* "to soothe pain", (14a), where the verb describes the alleviation of suffering. Finally, constructions are attested in which a cause appears as the subject and "pain" as the object, a pattern that allows the explicit representation of the agent responsible for increasing or reducing the sensation, as in *dolorem augere* "to increase pain" or *dolorem minuere* "to decrease pain", (14b):

(14a) Ovid, *Metamorphoses*, 13.317

…temptaretque feros requie ***lenire dolores***.

"and seek to soothe his terrible anguish by a time of rest."

(14b) Cicero, *Ad familiares*, 6.1.1

…tamen oculi ***augent dolorem****, qui ea quae ceteri audiunt intueri cogunt*.

"But the eyes add to the pain, forcing us to look at what others only hear."

4.2 Metaphors of *pain*

Many of the verbs in the constructions analysed carry a figurative meaning. As noted, pain intertwines physical, psychological, and cultural dimensions that are difficult to conceptualise and express. When describing this unpleasant experience, speakers often invoke causes of physical harm, even in cases of psychological pain (Semino 2010). Thus, images such as fire, sharp objects, or wild animals are frequently used, as they threaten bodily integrity and effectively convey intensity. These metaphors also aim to elicit an embodied, empathetic response in the listener or reader, highlighting the deeply corporeal grounding of pain expressions.

The close link between physical and mental pain (Fedriani 2025: 215) further explains this preference: psychological pain often somatises, while physical pain generates emotional distress. Embodied metaphors are therefore especially powerful, projecting familiar sensory experiences onto the mental domain and enabling us to understand and communicate suffering more effectively.

In this regard, one of the most salient features of metaphorical descriptions of pain is their high level of detail, creativity, and textual complexity (Semino 2010), traits that reflect the inherent difficulty of verbally representing such an intense and complex experience. Table 10.3 presents the metaphors in our corpus organised according to the type of SVC. [10]

10 All the source domains shown in Table 10.3 are attested in metaphors associated with emotions across many languages worldwide. Some of them are well-documented (Kövecses 2000: 37–38), such as "fire", "container", or "animal", whereas others, such as "fluid", have been proposed by Buccheri et al. (2021), see also Jiménez Martínez (2025).

SVC		Metaphor													Total
		MOVEMENT	LIVING BEING	HUMAN BEING	OPPONENT	WILD ANIMAL	OBJECT	BURDEN	CONTAINER	FLUID	FIRE	NATURAL FORCE	WOUND/ILLNESS	CHAIN	
General		2	3		1	1	4	12	3		6	1	2	1	36
Aspectual	Inchoative	10	13		17	1	7	3			3		1		55
	Terminative	1	1	1	1		1	5	3	2					15
Causative	Inchoative		1		1		14	9	1		2				28
	Terminative			2	1		1	10					1		15
Degree	Plus		6	3			3								12
	Minus			2			3								5
Total		13	24	8	21	2	33	39	7	2	11	1	4	1	166
Total %		7.8	14.5	4.8	12.7	1.2	19.9	23.5	4.2	1.2	6.6	0.6	2.4	0.6	100

Table 10.3 Distribution of the metaphors linked with *dolor* "pain" in our corpus of archaic and classical Latin.

4.2.1 General support-verb constructions: Pain is a burden and pain is fire

We begin with the metaphorical projections activated in general SVCs. From a quantitative perspective, it is evident that the most frequently employed source domain to express the experience of pain is PAIN is a BURDEN. In this way, the sensation of discomfort is conceptualised as a weight pressing down on the experiencer. The most representative SVC of this schema is *dolorem [per-, ab-]ferre* "to bear pain", (4a) and (15). However, this is not the only way in which this metaphor is realised. The authors under study also employ other constructions, such as *dolor grauet* "pain burdens", (16a) or *dolorem tolerare* "to tolerate pain", (16b), to convey the same underlying conceptual metaphor.

(15) Livy, *Ab urbe condita*, 8.4.8

pertulerunt *tamen* ***hunc dolorem****, bellum nostro nomine cum Paelignis gessimus.*

"Yet they swallowed their resentment. We waged war on our own account with the Peaeligni."

(16a) Seneca, *Phaedra*, 859

*effare aperte quis **grauet** mentem **dolor**.*

"Speak out openly: what pain is burdening her mind?"

(16b) Cicero, *In Verrem*, 2.5.118

*O magnum atque **intolerandum dolorem**!*

"Think of that unbearable burden of pain."

As Fedriani (2025: 224) notes, this source domain is structured into two sub-metaphors: one conceives that PAIN is a BURDEN THAT PRESSES DOWN ON THE EXPERIENCER, and the other focusses on the idea of LIBERATING FROM THE BURDEN. The first sub-metaphor is evident in the SVCs exemplified in (15) and (16), as well as in other constructions such as *dolor premit* "pain presses" or *dolor [re-, o-]primit* "pain represses/oppresses", which are inchoative aspectual in nature. The metaphor of release, in turn, appears in terminative aspectual combinations, such as *dolorem mittere* "to send (away) pain", (17) and *dolorem pellere* "to push (away) pain", and also in terminative causative SVCs, such as *dolorem leuare* "to alleviate pain", (18):

(17) Seneca, *Phoenissae*, 348-349

***Mitte** uiolentum impetum / **doloris**.*

"Check this violent outburst of your pain!"

(18) Cicero, *Ad familiares*, 6.4.2

*Hoc loco si uideor augere **dolorem tuum** quem consolando **leuare** debebam.*

"At this point, if you think I am aggravating your grief when I ought to be assuaging it."

Another important figurative schema to talk about pain is PAIN IS FIRE. As Kövecses (2000: 81–97) has amply demonstrated, the metaphor EMOTION IS HEAT or FIRE constitutes a fundamental conceptual resource for representing a wide range of human feelings whose most salient feature is intensity: anger, passion, love, desire, ambition, lust, hatred, or pain, among others.

In the case of pain, its association with fire and extreme heat can be explained in two ways. First, it reflects intense, almost unbearable emotional experiences, often merging with anger or related emotions. As Fedriani (2025: 230) observes, these metaphors commonly convey emotional suffering connected to anger, and the fire domain is central to Latin conceptualisations of anger, making this overlap significant. Second, the metaphors involve a metonymic component

grounded in bodily experience: fire causes actual physical pain. This causal link allows the fire domain to function both as a metaphor for emotional intensity and as a metonymic extension of physical sensation into the affective realm, integrating physical and emotional experience.

Our corpus records eleven SVCs based on this fire metaphor. To understand how Latin conceptualises pain through this metaphor, we will analyse the most productive constructions, distinguishing between: (a) those that express the internal experience of the emotion through intransitive verbs, formalising intensive extensions of general SVCs (six examples), and (b) those that metaphorically evoke the act of igniting the sensation in the experiencer through transitive verbs, that is, intensive variants of inchoative aspectual SVCs (three examples) and inchoative causative SVCs (two examples).

From the perspective of general SVCs, we find the metaphor SUFFERING PAIN is a FLAME THAT BURNS. This schema is realised in constructions such as *dolor[e] [ex]ardet* "pain burns", (19a), *dolor feruet* "pain blazes", (6b), and *dolor urit* "pain burns", (19b), in which the focus is on the idea that pain, like fire, is a dynamic and progressive reality that grows in size and intensity (cf. Roesch 2016; 2017).

(19a) Vergil, *Aeneid*, 9.66

haud aliter Rutulo muros et castra tuenti / ignescunt irae, ***duris dolor ossibus ardet***.

"just so, as he scans wall and camp, the Rutulian's wrath is aflame; resentment is hot within his iron bones."

(19b) Ovid, *Metamorphoses*, 4.257

tantus dolor urit *amantes*.

"so great is the burning smart which jealous lovers feel."

The conceptualisation of *dolor* as fire allows also the use of transitive verbs such as *incendere* "to kindle" (one example) and *accendere* "to set on fire" (two examples) to form constructions that emphasise the onset of the sensation. These constructions correspond to the metaphorical projection BEGINNING TO FEEL PAIN IS LIGHTING A FIRE, in which the activation of the emotional state is understood as the act of igniting a flame.

(20) Seneca, *Medea*, 671

immane quantum augescit et semet ***dolor*** */* ***accendit ipse*** *uimque praeteritam integrat.*

"Monstrously grows her grief, feeds its own fires and renews its former strength."

In the same vein of expressing the overwhelming intensity of a sensation through analogy with a natural force (PAIN is a NATURAL FORCE), the corpus attests the SVC *dolor exundat* "pain flows out", (21). In this construction, the experience of pain is depicted at an extreme level by comparing it to the violent force of waves during a storm.

(21) Seneca, *Oedipus*, 924

sudor per artus, spumat et uoluit minas / ac mersus alte magnus ***exundat dolor****.*

"limbs with cold sweat streaming, froths and threatens, and his mighty, deep-buried anguish overflows."

4.2.2 Inchoative support-verb constructions: pain is an opponent, pain is a living being and the beginning of an event is the endpoint of telic motion

As noted in Section 4.1.2, inchoative aspectual SVCs are the most frequent in the corpus. These constructions, referring to the spontaneous onset of pain, typically instantiate the metaphor PAIN is an OPPONENT (17 examples). Pain is thus conceptualised as an enemy that attacks the experiencer with intensity and violence: it can seize (*dolor rapit* "pain seizes"), overcome (*dolor [e]vincit* "pain overcomes"), suffocate (*dolor angitur* "pain is suffocated"), or disturb (*dolor turbat* "pain disturbs"). Through such verbs, the beginning of pain is framed as a forceful and often overwhelming assault.

(22a) Vergil, *Aeneid*, 4.474–475

Ergo ubi concepit furias ***evicta dolore*** */ decreuitque mori.*

"So when, conquered by anguish, she caught the madness and resolved to die."

(22b) Celsus, *De Medicina*, 2.7.25

Auris quoque ***dolor*** *acutus cum febre continua uehementique saepe mentem* ***turbat****.*

"Acute pain in ear with continuous severe fever also often disturbs the mind."

The opponent metaphor is typically associated with the beginning of negative emotions, such as anger, hatred, or rage, which involve experiences of violence, aggression, or helplessness. Indeed, this image is one of the most productive for describing emotional states that affect human beings (Jiménez Martínez 2025; Buccheri *et al.*, 2021: 170). Closely related is the metaphor PAIN is a WILD ANIMAL, which portrays pain as attacking with uncontrolled strength, (6a). This vivid, concrete imagery renders an otherwise intangible sensation more expressible.

The second most productive metaphorical domain among inchoative

aspectual SVCs featuring the noun *dolor* can be formulated as THE BEGINNING OF PAIN IS THE BIRTH OF A LIVING BEING. In this framework, coming into the world becomes a figurative way of referring to the beginning of an abstract event. Within this pattern, we find constructions such as *dolor oritur* "pain arises" and *dolor nascitur* "pain is born":

(23a) Plautus, *Stichus*, 165

uteri ***dolores*** *mihi* ***oboriuntur*** *cotidie.*

"I get my pains there every day."

(23b) Celsus, *De Medicina*, 1pr.23.1

Praeter haec, cum in interioribus partibus et ***dolores*** *et morborum uaria genera* ***nascantur****…*

"Moreover, as pains, and also various kinds of diseases, arise in the more internal parts…"

This figurative domain extends beyond inchoative SVCs. Particularly productive are degree constructions in which the increase in the intensity of pain is expressed through the image of a living organism that grows, as *dolor [in]crescit* "pain grows" in (4e).

Another highly productive metaphorical schema for expressing the onset of pain is the one articulated through verbs of movement and spatial orientation, such as *dolor peruenit* "pain arrives" or *dolor transit* "pain passes", already examined in (Section 4.1) and in example (9). In this group of constructions, known as orientational metaphors (Fedriani 2025; Jiménez Martínez 2020), pain is conceptualised as an entity that moves, advances, settles, or passes through the experiencer's body.

Unlike the metaphors discussed above, which render the internal experience of pain as objects, animals, fire, or even persons, these metaphors activate a spatial projection: pain behaves as a mobile element within a territory, the body. In this way, the different stages of the painful episode are metaphorically expressed through patterns of movement and position in physical space, drawing on the very general and schematic macro-metaphor EVENTS are MOVEMENTS (Lakoff and Johnson 1980: 25).

(24a) Cicero, *Ad familiares*, 14.9.1

Ad ceteras meas miserias ***accessit dolor*** *et de Dollabellae ualetudine et de Tullia.*

"To all my other woes must now be added the sorrow caused me by the ill-health of Dolabella and Tullia."

(24b) Celsus, *De medicina*, 1pr.25.1

*neque enim, cum **dolor** intus **incidit**, scire quid doleat eum.*

"For when pain occurs internally, neither is it possible for one to learn what hurts the patient."

In example (24a), pain is construed as a material entity entering the speaker's emotional sphere. The verb *accedere* "to approach" denotes telic motion, oriented toward an endpoint that, when abstract, marks the beginning of the situation, process, or event represented by the noun, following the metaphor THE BEGINNING OF AN EVENT IS THE ENDPOINT OF A TELIC MOTION. Similarly, *incidit* "to fall" in (24b) conveys the sudden onset of pain rather than physical falling. As a telic verb, it implies lack of control and agency, often associated with negative events (*in vitium incidere* "fault strikes", *in morbum incidere* "to fall ill"), activating the metaphor THE BEGINNING OF A NEGATIVE SITUATION IS A FALL. This reflects the general schema DOWN IS BAD and introduces an intensity nuance absent in *dolor accedit* "pain approaches".

4.2.3 Causative support-verb constructions: pain is an object

We continue our analysis of SVCs with the noun *dolor* by examining the inchoative causative constructions. In these cases, one metaphorical schema clearly predominates: pain is an object (15 instances). More specifically, the most frequent construction in our corpus—*dolore adficere* "to inflict pain" (10 examples)—conceptualises this object as a tool or weapon capable of inflicting harm, that is, of producing the painful experience, (25).

(25) Cicero, *In Verrem*, 2.4.83

*hic uero tanto **dolore adficior** ut nihil mihi indignius, nihil minus ferendum esse uideatur.*

"Here, indeed, I am so overwhelmed by pain that nothing seems to me more unworthy or less bearable."

When this sensation is conceptualised as a thing, it is understood as an entity that can be produced, (26), possessed, (27a), taken up, or thrown, (27b):

(26) Celsus, *De medicina*, 2.3.6

*Si inflatio in superioribus partibus **dolorem** tumoremque **fecit**, bonum signum est sonus uentris.*

"When flatulence causes pain and swelling in the upper part of the abdomen, it is a good sign when intestinal rumbling passes thence downwards towards the lower belly."

(27a) Celsus, *De medicina*, 1.5.2

tum iis praecipue, qui uel capitis uel arteriae ***dolores habere*** *consuerunt, uel quoslibet alios oris adfectus*.

"it is especially so to those who are disposed to pains in the head or windpipe, or to other forms of oral affections."

(27b) Cicero, *Ad familiares*, 5.13.3

…quamquam tuis monitis praeceptisque omnis est ***abiciendus dolor***.

"and yet, according to your admonitions and instructions, all sorrow must be thrown to the winds."

In (26), *dolorem facere* "to cause pain" conveys the idea of generating pain in someone.[11] In this case, the inflammation of the upper part of the body produces pain in the patient. It is therefore a causative incoative SVC. However, this metaphor appears across almost all semantic types of SVCs with *dolor*, since the idea that EVENTS are OBJECTS constitutes a widespread metaphorical domain through which we communicate in everyday language across all languages. Thanks to this schema, we are able to speak about the beginning, duration, or end of any kind of abstract situation. Thus, for example, the construction *dolorem abicere* "to cast away pain", (27b), by incorporating a verb that denotes an action in which the possessed object is forcefully released, can be interpreted as conveying a terminative and intensive meaning.

A particularly interesting specification of this metaphor is PAIN is a CHAIN, as shown in (28). Through this image, Cicero exploits the physical connotations of the object—its capacity to bind, restrain, and immobilise—in order to express how pain can hold us captive and render us motionless:

11 Special attention must be paid to the verbs *habere* "to have" and *facere* "to do, to make" in the theory of SVCs. As Gross (2022: 3) explains, "la conjugaison des prédicats nominaux est dépendante de la nature sémantique des prédicats nominaux, ce qui n'est pas le cas des verbes. Ces supports sont fonction, dans une première approximation, de grandes classes, comme les actions (*faire*), les états (*avoir*) et les événements (*avoir lieu*)." (translation: "the conjugation of nominal predicates depends on the semantic type of the nominal predicates, which is not the case for verbs. These support (sc. verbs) are, firstly, divided into broad groups, such as actions (*faire* "to do"), states (*avoir* "to have"), and events (*avoir lieu* "to take place").") In SVCs with *dolor*, a state noun, *dolorem habere* "to have pain" represents a prototypical construction in which the verb functions as a structural support expressing the experience of the sensation. Likewise, *facere* "to make" operates as a general causative verb, encoding the broad notion of "causing". Their inclusion in the metaphorical analysis is justified by the figurative interpretation of constructions with *dolor*: although both are prototypical support verbs, they convey abstract possession and causation, that is, metaphorical meanings.

(28) Cicero, *Ad familiares*, 7.1.1

Si ***te dolor*** *aliqui corporis aut infirmitas ualetudinis tuae* ***tenuit*** *quo minus ad ludos uenires*.

"If it was some bodily pain or weakness of health that prevented you coming to the games."

4.2.4 Terminative support-verb constructions: pain is a fluid and pain is a container

The final group of metaphors we will analyse is primarily associated with terminative SVCs, that is, constructions that predicate the final moment of the pain suffered by the experiencer. To convey this, the authors in our corpus, in addition to the analogy of a burden from which one must be freed, draw on two further metaphorical strategies. On the one hand, they create an analogy between pain and a fluid (PAIN is a FLUID), which can dissipate just as a drink disappears when someone quenches their thirst, as illustrated in the following example from Cicero (first century BCE):

(29) Cicero, *Ad familiares*, 5.16.4

non facile ***exhauriri*** *tibi istum* ***dolorem*** *posse uniuersum puto*.

"I doubt if your mind can easily be purged of so personal a sorrow altogether."

On the other hand, the authors link pain to a container that can be filled or emptied (PAIN is a CONTAINER). In *dolorem saturare* "to satiate pain" (30), pain is understood as a vessel that is not yet full, which evokes a further image often associated with this verb: that of a wild animal whose hunger is not easily sated. The combination of these two schemas produces a particularly vivid image of the cessation of pain:

(30) Vergil, *Aeneid*, 5.608

multa mouens necdum antiquum ***saturata dolorem***.

"Moving many things, not yet satisfied with ancient pain."

5. Intensity in the expression of *pain* in Latin

As Pérez Rull (2002: 324) observes, emotional intensity is an abstract notion that can only be grasped through more concrete concepts such as depth, quantity, force, or impact (also in Kövecses 1986: 82; 2000: 41). In our archaic and classical Latin corpus with *dolor*, this intensity is expressed through two main mechanisms.

First, degree-type SVCs depict pain as increasing or decreasing along a scale that quantifies its magnitude. When pain "grows" (*dolor crescit* "pain grows", *dolorem augere* "to increase pain"), it is felt as more intense; when it "lessens" (*dolorem minuere* "to reduce pain"), it is weaker. This mechanism relies on metaphors that conceptualise intensity as size or extension (EMOTIONAL INTENSITY IS SIZE).

The second, and more relevant type for our purposes, involves intensive extensions of support verbs in SVCs. These constructions conceptualise pain through metaphors drawn from domains of strength, impact, or violence, presenting it as forceful, sudden, or overwhelming. In this respect, the linguistic expression of pain is particularly revealing. As a universally graded experience, both physical and emotional pain is frequently conceptualised in terms of degree. Although some metaphors with *dolor* function as cognitive frameworks that help us navigate its concept, most of the metaphorical SVCs identified with this noun directly or indirectly encode the intensity of the experience.

Overall, we can say that pain intensity in archaic and classical Latin is expressed through the association of this sensation with violent everyday experiences (PAIN IS AN OPPONENT, PAIN IS A DAMAGING OBJECT, PAIN IS A WILD ANIMAL, PAIN IS FIRE, OR PAIN IS A DESTRUCTIVE NATURAL FORCE) or with very heavy material entities that exhaust, suffocate, paralyze, or immobilise us, such as PAIN IS A BURDEN OR PAIN IS A CHAIN.

6. Some remarks on diachrony: *dolor* in late Latin and medieval Spanish

We now take a temporal leap to examine the evolution of SVCs with the noun *dolor* in late Latin and medieval Spanish. This shift allows us to address a key linguistic question: how did Latin function in the Christian era with respect to expressing pain, and to what extent did early Spanish, as Latin's direct heir, inherit these figurative patterns? Specifically, we ask whether late Latin and medieval Spanish preserved the rich metaphorical repertoire of classical Latin or whether they developed new metaphorical projections shaped by their own linguistic and cultural contexts.

It is important to note the limitations of the corpus, especially for the Romance language: surviving texts rarely reflect everyday speech, as they mostly belong to legal, administrative, or formulaic genres that often retain archaism or clerical conventions. Nonetheless, our aim is to outline the main features of pain expressions in this period and provide a preliminary diachronic reflection based on the available evidence. We begin by presenting our late Latin and medieval Spanish data in Table 10.4.

	Late Latin		Medieval Spanish			
Types of SVC	SVC	Tokens	Total	SVC	Tokens	Total
General	*dolor cruciat* “pain crucifies”	14	85	Exp. *ha dolor* “sb. has pain”	113	197
	aliquis Exp. *dolorem patitur* “sb. suffers pain”	9		Exp. *tiene dolor* “sb. holds/is in pain”	12	
	aliquis Exp. *dolore ardet* “sb. burns with pain”	8		Exp. *sufre dolor* “sb. suffers pain”	12	
	dolor (at)terit “pain rubs”	8		Exp. *siente dolor* “sb. feels pain”	12	
	aliquis Exp. *dolorem sentit* “sb. experiences pain”	6		Exp. *dolor constriñe al* “pain constrains sb.”	10	
	dolor torquet “pain twists”	5		Exp. *dolor atormenta al* “pain plagues sb.”	7	
	aliquis Exp. *plenus est dolore* “sb. is full of pain”	5		Exp. *pasa dolor* “sb. is in pain”	6	
	dolor consumit “pain consumes”	5		Exp. *dolor turba al* “pain disturbs sb.”	5	
	aliquis Exp. *dolorem (con / su)fert* “sb. undergoes pain”	5		Exp. *está en el dolor* “sb. is in pain”	4	
	aliquis Exp. *dolorem habet* “sb. has/is in pain”	4		Exp. *dolor queda en el* “pain remains in sb.”	3	
	aliquis Exp *dolorem / in dolore sustinet* “sb. remains in pain”	4		Exp. *dolor embarga al* “pain overwhelms sb.”	2	
	dolor manet “pain remains”	2		Exp. *vive en el dolor* “sb. lives in pain”	2	
	aliquis Exp. *dolorem retinet intus* “sb. holds pain inside”	2		Exp. *padece dolor* “sb. suffers from pain”	2	
	aliquis Exp. *dolorem portat* “sb. carries pain”	2		Others	7	
	aliquis Exp. *dolorem tolerat* “sb. tolerates pain”	2				
	Others	4				

Aspectual	Inchoative	*dolor percellit* “pain beats down”	6	57	*dolor vence* “pain beats”	4	28
		dolor inuenit “pain arrives”	5		Exp. *toma el dolor* “sb. receives pain”	3	
		aliquis Exp. *(sus / con)cipit dolorem* “sb. receives pain”	5		Exp. *dolor toma al* “pain grabs sb.”	3	
		dolor mordet “pain bites”	3		Exp. *dolor prende al* “pain holds on to sb.”	3	
		dolor corrumpit “pain destroys”	3		Exp. *dolor viene al* “pain comes to sb.”	2	
		dolor stimulat “pain incites”	3		Exp. *dolor hiere al* “pain hurts sb.”	2	
		dolor affligit “pain casts down”	3		Exp. *engendra el dolor* “sb. produces pain”	2	
		aliquis Exp. *parit dolorem* “sb. produces pain”	3		Exp. *se mueve al dolor* “sb. moves to pain”	2	
		dolor mouit “pain moves”	3		Others	7	
		dolor pellit “pain strikes”	2				
		dolor tangit “pain touches”	2				
		dolor (ir / ob)ruit “pain tumbles”	2				
		Others	17				
	Terminative	*dolor fugit* “pain flees”	3	15	Exp. *huye del dolor* “sb. runs away from pain”	1	3
		dolor ad salutem aperitur “pain is taken away for wellbeing”	3		*dolor huye del* Exp. “pain runs away from sb.”	1	
		aliquis Exp. *ab dolore fugit* “sb. flees from pain”	2		*dolor parte del* Exp. “pain departs from sb.”	1	
		aliquis Exp. *liber(atus) ab dolore est* “sb. is liberated from pain”	2				
		Others	5				

Causative	Inchoative	*aliquis* Caus. *aliquem* Exp. *dolore adficit* “sb. inflicts pain on sb.”	3	20	Caus. *da dolor al* Exp. “sb. causes pain to sb.”	4	14
		aliquis Caus. *aliquem* Exp. *in dolorem auertit* “sb. removes sb. from pain”	2		Caus. *hace dolor al* Exp. “sb. puts pain on sb.”	3	
		aliquis Caus. *aliquem* Exp. *in dolore point* “sb. puts sb. in pain”	2		Caus. *trae dolor al* Exp. “sb. causes pain to sb.”	3	
		aliquis Caus. *dolorem generat* “sb. genereates pain”	2		Caus. *causa dolor al* Exp. “sb. causes pain to sb.”	3	
		aliquis Caus. *dolorem ingerit* “sb. heaps pain upon”	2		Others	1	
		Others	9				
	Terminative	*aliquis* Caus. *dolorem sanat* “sb. cures pain”	5	14	Caus. *medicina el dolor* “sb. manages pain”	8	28
		aliquis Caus. *dolorem curat* “sb. cares for pain”	4		Caus. *tira el dolor* “sb. eases pain”	6	
		Others	5		Caus. *tolle el dolor* “sb. takes away pain”	5	
					Caus. *sana el dolor* “sb. soothes pain”	4	
					Caus. *pone remedio al dolor* “sb. gives a remedy for pain”	2	
					Others	3	
Degree	Plus	*aliquis* Caus. *dolorem geminat* “sb. doubles pain”	2	8	Caus. *acrecienta el dolor* “sb. makes pain worse”	3	9
		aliquis Caus. *dolorem multiplicat* “sb. multiplies pain”	2		*dolor crece* “pain grows”	3	
		Others	4		Others	3	
	Minus	Others	2	2	Caus. *amansa el dolor* “sb. soothes the pain”	7	14
					dolor mengua “pain subsides”	5	
					Others	2	
Total		76 SVC types		201	60 SVC types		293

Table 10.4 Distribution of SVCs with *dolor* “pain” in late Latin and medieval Spanish.

The most immediate conclusion drawn from the table is that expressions of pain in late Latin and medieval Spanish display considerably less diversity than in archaic and classical Latin. In the earlier period, 166 examples are distributed across 104 constructions; in late Latin, 201 examples occur in 76 SVCs; and in medieval Spanish, 293 examples are concentrated in only 60 constructions. This distribution suggests a gradual reduction in creative variation after the classical period.[12]

Beyond this quantitative trend, the overall picture reveals a clear diachronic development. Late Latin occupies an intermediate position between archaic and classical Latin and medieval Spanish. Like the earlier stage, seventh- to ninth-century Latin still assigns an important role to inchoative SVCs—both aspectual and causative—in conceptualising pain, highlighting the cultural salience of its emergence. At the same time, however, late Latin grants greater prominence to general constructions referring to the experience of pain itself (85 examples), a tendency that becomes dominant in medieval Spanish. There, *haver dolor* "to have pain" clearly predominates with 113 occurrences (39% of the total), whereas in late Latin no construction exceeds 15 tokens, reflecting a more even distribution.

The remaining SVCs occur with much lower frequencies in both periods: even relatively general constructions such as *dolorem pati* "to suffer pain" and *dolore ardere* "to burn with pain" in late Latin, or *tener dolor* "to have pain" and *sentir dolor* "to experience pain" in medieval Spanish, reach no more than 12 examples each. This quantitative imbalance suggests not only reduced lexical productivity but also a gradual narrowing of metaphorical variation. While only moderately perceptible in late Latin, this tendency becomes markedly pronounced in medieval Spanish, where the dominance of a single construction points to an ongoing process of conventionalisation in the expression of pain.

Regarding the metaphorical expression of pain, late Latin displays a distribution largely comparable to that observed in archaic and classical Latin. Medieval Spanish, by contrast, exhibits a noticeably narrower range of figurative patterns than Latin in either period. The data thus suggest continuity between the two Latin stages, alongside a subsequent restriction in metaphorical diversity in medieval Spanish. Below, we present the metaphorical schemas identified in the corpus, organised according to the semantic types of SVCs (see Table 10.5).

12 Cf. n. 5 for the limitations of the data collection.

SVC			Metaphor													Total
			MOVEMENT	LIVING BEING	HUMAN BEING	OPPONENT	WILD ANIMAL	OBJECT	BURDEN	CONTAINER	FLUID	FIRE	NATURAL FORCE	WOUND/ILLNESS	CHAIN	
LATE LATIN	General		4			27		10	11	3	6	13		9	2	85
	Aspectual	Inchoative	7	8		27	5	7		1	1		1			57
		Terminative	2	2	1	5		3							2	15
	Causative	Inchoative	1	3		1		10		4		1				20
		Terminative				1		1	2		1			9		14
	Degree	Plus		2				5	1							8
		Minus						1				1				2
	Total		14	15	1	61	5	37	14	8	8	15	1	18	4	201
	Total %		7	7.4	0.5	30.3	2.5	18.4	7	4	4	7.4	0.5	9	2	100
MEDIEVAL SPANISH	General					26		129	6	6				27	3	197
	Aspectual	Inchoative	5	11		6		3		3						28
		Terminative		1		2										3
	Causative	Inchoative						14								14
		Terminative						12	1					15		28
	Degree	Plus		4				5								9
		Minus		1			7	5					1			14
	Total		5	17		34	7	168	7	9			1	42	3	293
	Total %		1.7	5.8		11.6	2.4	57.4	2.4	3.1			0.3	14.3	1	100

Table 10.5 Distribution of the metaphors linked with *dolor* "pain" in our corpus of late Latin and medieval Spanish.

A review of Table 10.5 reveals several key findings. Regarding late Latin, its metaphorical landscape resembles that of archaic and classical Latin, displaying comparable richness and diversity, but with some novel tendencies—such as the predominance of the metaphor PAIN is an OPPONENT (61 examples)—and a decline in the prominence of others, like PAIN is a BURDEN. In contrast, medieval Spanish appears less creative than archaic and classical Latin in expressing pain and tends to conceptualise it more generally, often through the image of an object, a pattern consistent across all semantic types of SVC.

Concerning the situation of medieval Spanish, it is striking that no examples of

SVC based on the metaphor of fire—so prominent in Latin—appear in our corpus, and that there are only six instances of the PAIN is a BURDEN schema. Additionally, the metaphors PAIN is a FLUID and PAIN is a HUMAN BEING are completely absent, suggesting a simplification or selective adoption of metaphorical projections.

As for the second observation, the systematic use of the metaphorical pattern PAIN is an OBJECT, a simple and generalised projection, reinforces the impression that medieval Spanish, at least in the texts studied, lacks the metaphorical richness found in Latin of both moments. Pain is thus conceptualised primarily in terms of the experience itself and through accessible cognitive schemas, whereas Latin displays greater variety and creativity in representing its different dimensions, at least in the contexts we have examined. To illustrate this, we present Figure 10.1, which clearly reflects the preferences for metaphors in each language.

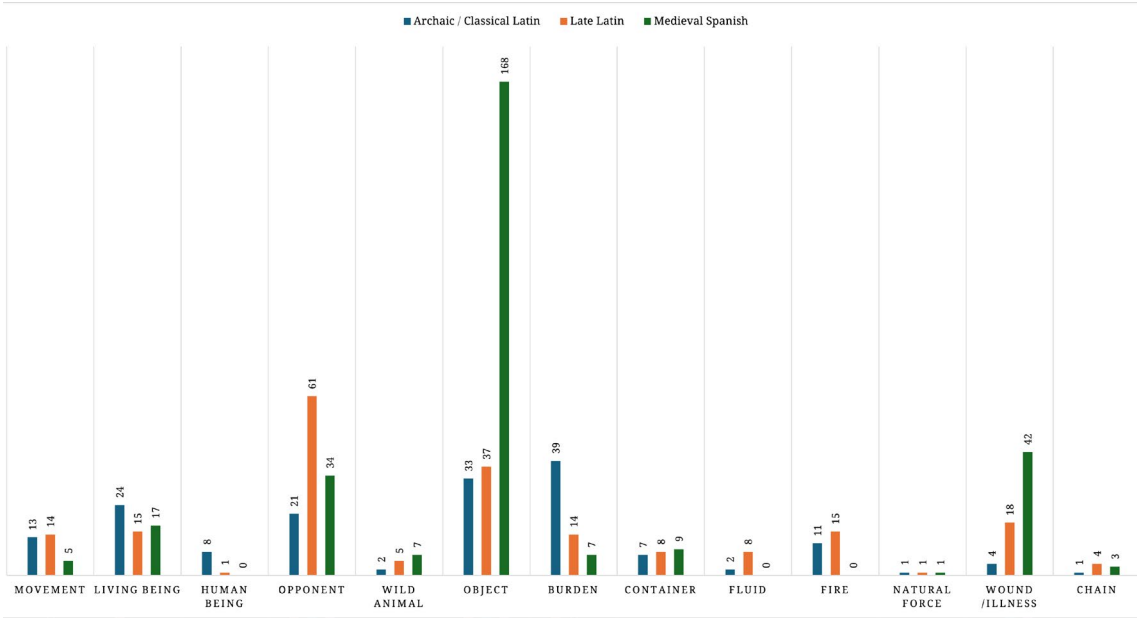


Fig. 10.1 Distribution of the metaphors of *dolor* "pain" in the periods studied. Figure created by the author (2026).

Although the lower diversity of metaphorical schemas in medieval Spanish reflects accurately the corpus data, these findings must be interpreted cautiously, as emphasised throughout this study. The limited variety of Spanish textual genres likely prevents us from detecting many metaphorical patterns that consistently appeared in Latin. It is therefore plausible that metaphors present in late Latin—a language spoken in the same regions only two centuries earlier—persisted into thirteenth-century Spanish. Deeply ingrained cognitive schemas are unlikely to disappear within a few generations. The most plausible explanation is thus either (i) that surviving texts fail to capture these structures,

perhaps preserved in oral genres or lost over time, or (ii) that these metaphors lived on in constructions excluded from this study, such as SVCs without the noun *pain* or in idiomatic expressions.

Finally, the expression of the intensity inherent to this sensation—although our data of medieval Spanish relies less on SVCs than Latin—is mainly conveyed through the same strategies as in archaic and classical Latin: degree constructions, such as *dolor ingrauescit* "pain grows burdensome" or *amansar el dolor* "to ease the pain", and intensified extensions of support verbs. Two strategies stand out: the notion of violence, reflected in the metaphor PAIN is an OPPONENT, and the concept of a material force that immobilises, as in PAIN is a CHAIN. In late Latin, the association of pain with fire (PAIN is FIRE) is also particularly prominent.

7. Conclusions

Throughout this corpus-based study, in which we analysed the expression of pain through SVCs in archaic and classical Latin (second century BCE to the second century CE), late Latin (seventh to ninth centuries CE), and medieval Spanish (thirteenth to fifteenth centuries CE), we observed that constructions expressing the experience of pain—characterised by an inherent intensity—span a wide range of metaphorical patterns, covering all stages and dimensions of this experience.[13]

In Latin, we identified an extensive collocational network rich in metaphorical projections that describe, with considerable creativity and detail, an experience notoriously difficult to express in words. Two main strategies stand out for conveying the intensity of pain: first, the notion of violence that assaults and terrifies the experiencer, expressed through metaphors such as *pain is an opponent, a wild animal, a weapon*, or *fire*; and second, the conception of pain as an extremely heavy material reality that immobilises the sufferer, reflecting both physical and mental aspects of the experience.

Examining late Latin reveals a similarly rich, diverse, and creative set of images used to express pain in the seventh to ninth centuries. The data from medieval Spanish, however, show a notable reduction in complexity and diversity in generating pain metaphors. Most SVCs are based on the simple figurative schema *pain is an object*, likely reflecting the more limited range of textual genres studied rather than the actual linguistic reality, in which the same cognitive schemas probably persisted as in Latin three centuries earlier. Nevertheless, even with less elaborate examples, medieval Spanish employs the same mechanisms as Latin—violence and immobilisation—to convey the intensity of pain, suggesting the persistence of a generalised cognitive pattern in the conceptualisation of this experience, as noted by Kövecses (2006).

13 Cf. n. 5 for the limitations of the data collection.

Acknowledgements

This work is part of the research project *Interacción del léxico y la sintaxis en griego antiguo y latín 2: Diccionario de Colocaciones Latinas (DiCoLat) y Diccionario de Colocaciones del Griego Antiguo (DiCoGra)* (PID 2021-125076NB-C42).

Bibliography

Baños, José Miguel (ed.). 2023. *Sintaxis latina. Volumen II: Sintaxis de la oración compleja: Orden de palabras*. Madrid: Manuales y anejos de Emerita LVI, Consejo Superior de Investigaciones Científicas.

Baños, José Miguel, and María Dolores Jiménez López. 2024. Support-verb Constructions in the Gospels: A Comparative Study between Greek and Latin. In Victoria Beatrix Fendel (ed.), *Support Verb Constructions in the Corpora of Greek: Between Lexicon and Grammar?*, 93–129. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.14017927

Baños, José Miguel, María Dolores Jiménez López, María Isabel Jiménez Martínez, and Cristina Tur (eds.). 2022. *Collocations in Theoretical and Applied Linguistics: From Classical to Romance Languages*. Madrid: Guillermo Escolar Editor.

Bodelot, Colette, and Olga Spevak (eds.). 2018. *Les constructions à verbe support en latin*. Clermont-Ferrand: Presses Universitaires Blaise Pascal.

Buccheri, Alessandro, Irene De Felice, Chiara Fedriani, and William M. Short. 2021. Semantic Analysis and Frequency Effects of Conceptual Metaphors of Emotions in Latin. From a Corpus-based Approach to a Dictionary of Latin Metaphors. *Journal of Latin Linguistics* 20. 163–189. https://doi.org/10.1515/joll-2021-2002

Cicalese, Anna. 1999. Le estensioni di verbo supporto: Uno studio introduttivo. *Studi Italiani di Linguistica Teorica e Applicata* 3. 447–485.

Deignan, Alice. 2005. *Metaphor and Corpus linguistics*. Amsterdam and Philadelphia, PN: John Benjamins. https://doi.org/10.1075/celcr.6

Fedriani, Chiara, and Irene De Felice. 2022. Collocazioni verticali: metafore di orientamento UP/DOWN nella lingua Latina. In José Miguel Baños et al. (eds.), *Collocations in Theoretical and Applied Linguistics: From Classical to Romance Languages*, 81–120. Madrid: SEEC-Guillermo Escolar Editor.

Fedriani, Chiara. 2025. Dolor Metaphors in Latin: What a Corpus-Based Approach to Ancient Sources Can (and Cannot) Tell Us. In Fabian Horn (ed.), *Metaphors of the Ancient World: Conceptual Metaphor Theory and Ancient Sources*, 215–243. Berlin and Boston, MA: Mouton de Gruyter. https://doi.org/10.1515/9783111632704-009

Gavriilidou, Zoé. 2004. Verbes supports et intensité en grec modern. *Lingvisticæ Investigationes* 27(2). 295–308. https://doi.org/10.1075/li.27.2.11gav

Gross, Gaston. 2022. Pour un recensement systématique des verbes supports. *Neophilologica* 31. 1–21. https://doi.org/10.31261/neo.2022.34.02

Jiménez Martínez, María Isabel. 2020. El movimiento como metáfora. Colocaciones con los verbos de movimiento venio e incido en latín. *Nova Tellus* 38(2). 177-196. https://doi.org/10.19130/iifl.nt.2020.38.2.0008

— 2025. Cupiditate ardere: estudio de las metáforas latinas del deseo. *Revue de Linguistique Latine du Centre Alfred Ernout (De Lingua Latina)*. 1–42. https://doi.org/10.62806/aamt3098

Kövecses, Zoltán. 2000. *Metaphor and Emotion: Language, Culture and Body in Human Feeling.* New York: Cambridge University Press.

— 2002. *Metaphor: A Practical Introduction*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780195145113.001.0001

Lakoff, George. 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind.* Chicago, IL: University of Chicago Press.

Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live by*. Chicago, IL: University of Chicago Press.

Langacker, Ronald. 1987. *Foundations of Cognitive Grammar. Vol. I. Theoretical Prerequisites.* Stanford, CA: Stanford University Press.

— 1991. *Foundations of Cognitive Grammar. Vol. II. Descriptive Application.* California, CA: Stanford University Press.

Leotta, Roberta G. 2023. Phenomenon of Jealousy in Latin Language and Literature: Metaphors, Scenarios and Embodied Experiences. PhD thesis (University of Edinburgh).

Malik, Norasyikin Abdul, Faizah Mahamad, and Mahamad Syafiq Ya Shak. 2025. Understanding Metaphor Collocation and its Patterns. *International Journal of Research and Innovation in Social Science* 9(7). 5416–5424. https://doi.org/10.47772/ijriss.2025.907000437

Mendózar Cruz, Juan. 2019. La expresión de la causatividad en latín: diátesis léxica y colocaciones. PhD thesis (Universidad Complutense de Madrid).

— 2020. Las construcciones con verbo soporte en latín. Estado de la cuestión. *Tempus* 47. 7-48.

Patekar, Jakob. 2022. What is a Metaphorical Collocation? *Fluminensia*, 34(1). 31–49. https://doi.org/10.31820/f.34.1.5

Roesch, Sophie. 2016. Le feu de la colère en latin: une métaphore basée sur la temperature. In Paolo Poccetti (ed.), *Latinitatis rationes. Descriptive and Historical Accounts for the Latin Language*, 445–463. Berlin and Boston, MA: Mouton de Gruyter. https://doi.org/10.1515/9783110431896-030

— 2017. Incendere: “incendier”, “augmenter” ou “remplir”? Étude sémantique. *Pallas* 103. 191–200. https://doi.org/10.4000/pallas.4163

Salas, Guillermo. 2024. Impersonal Inchoative Support Verb Constructions with ineo: Syntax, Semantics and Conceptual Structure. In Irene De Felice and Chiara Fedriani (eds.), *Studies in Cognitive Classical Semantics*, 37–52. Alessandria: Edizioni dell’Orso.

Semino, Elena. 2010. Descriptions of Pain, Metaphor, and Embodied Simulation. *Metaphor and Symbol* 25. 205–226. https://doi.org/10.1080/10926488.2010.510926

Stefanowitsch, Anatol. 2006. Corpus-based Approaches to Metaphor and Metonymy. In Anatol Stefanowitsch and Stefan T. Gries (eds.), *Corpus-based Approaches to Metaphor and Metonymy*, 1–16. Berlin and New York: Mouton de Gruyter. https://doi.org/10.1515/9783110199895.1

11. Epilogue: A brave new world beyond the canon

Victoria Beatrix Fendel

According to the *Oxford English Dictionary* (Mcarthur et al. 2018), a canon is a general rule, a standard for judgement or authority, or a baseline epoch or era.

> Oxford English Dictionary, s. v. canon (noun)
>
> …
>
> 2. Gen[eral]
>
> 2a. A law, rule, edict (other than ecclesiastical) [already in Old English]
>
> 2b. A general rule, fundamental principle, aphorism, or axiom governing the systematic or scientific treatment of a subject; e.g. canons of descent or inheritance; a logical, grammatical, or metrical canon; canons of criticism, taste, art, etc. [since 1588]
>
> 2c. A standard of judgement or authority; a test, criterion, means of discrimination [since 1601]
>
> …
>
> 4. The collection or list of books of the Bible accepted by the Christian Church as genuine and inspired. Also transferred, any set of sacred books; also, those writings of a secular author accepted as authentic [since 1382]
>
> …
>
> 9a. A chief epoch or era, serving to date from (Greek κανὼν χρονικός); a basis for chronology [since 1833]

The existence of such a canon—here, a canon of texts—presupposes a wider body of witnesses, in our case textual ones, that fall outside of the canon because they do not conform to its governing rule(s) (see 2b), its standard(s) of judgement and authority (see 2c), or its temporal limits in the sense of a chief epoch or era (see 9a) (see especially Chapter 4, which questions the concept of a classical period, and Chapter 8 which calls the canon of the New Testament "artificial"). Throughout this volume, we have focussed on enhancing the visibility of non-canonical

©2026 Victoria Beatrix Fendel, CC BY-NC 4.0 https://doi.org/10.11647/OBP.0478.11

datasets, both physical and virtual ones, by showcasing methodologies to do so (see especially Part I), on promoting the interoperability of methodological approaches, through transparent definitions and guidelines (see especially Part II), and on furthering the replicability of analyses, through empirically founded approaches with interoperable datasets(see especially Part III).

Verbal multi-word expressions are not a type of structure that is limited to any one register, genre, dialect, sociolect, or idiolect in any of the language varieties assessed here (unlike Marušić 2018 for German). Nor do all verbal multi-verb expressions qualify as "idioms" or "phrases" as often assumed for corpus languages (see further Fendel 2024). They are, in fact, everywhere, a natural part of the lexicon and the grammar. We have observed them throughout this volume in funerary inscriptions from Old Kingdom pyramids (Chapter 2), philosophical papyri from the Herculanean school (Chapter 3), medical treatises associated with the doctors Hippocrates and Galen (Chapter 4), Byzantine saints' lives (Chapter 6), biblical texts of the Old and New Testaments and their translation into various languages (Chapters 8 and 9), the writings of Coptic abbots (especially Shenoute of Atripe, in the *Coptic Scriptorium* and Leipoldt (1913) (Chapter 7), and Latin/Romance verse literature across centuries (Chapter 10).

Corpora that are less visible for reasons of materiality (for example, carbonised papyrus scrolls (Figure 11.1), inscriptions battered by wind and rain, e.g. ISic000136 = CIL X.7379,[1] or papyri exposed to oxygen and a range of little animals),[2] subject matter (for example, medical procedures and anatomical observations), research tradition (for example, the disciplinary affiliation of biblical texts and Byzantine saints' lives), etc. have much to offer.



Fig. 11.1 Herculaneum Papyrus number 1251, cornice 1 curated at the Officina dei Papiri in the Biblioteca Nazionale "Vittorio Emanuele III" di Napoli; traditionally identified as Philodemus' treatise *On Choices* (Περὶ Αἱρέσεων καὶ Φυγῶν *Peri Haireseōn kai Phugōn*). The image is courtesy of Brigham Young University, Provo, Utah and the Ministero per i Beni Culturali. All rights reserved.

1 ISic000136, from Termini Imerese, Sicily, dated to the first to third centuries CE, funerary inscription, Latin: https://isicily.classics.ox.ac.uk/inscription/ISic000136

2 P. Herm. 7, from Lykopolis, Egypt, dated to the fourth century CE, personal letter, Greek: https://luna.manchester.ac.uk/luna/servlet/view/search?q=metadata_schema=21126

They compare and contrast with our canonical texts, thus shining a new light on both. They can be assessed and made more visible by means of the tools shown in this volume, including Maque-IT (see D'Angelo and Nicolardi 2024), the building of Sketch Engine corpora, and treebanking. We hope that by making all datasets and codebases openly accessible, we can inspire further research into the corpora presented here. By making our definitions, frameworks, and methodological processes transparent, we hope to contribute to this further. A further step to take in order to ensure replicability and transparent reporting of all results would be pre-registration of datasets (for example, Hervais-Adelman et al. 2025: 311; Santiago et al. 2025: 342–343). A further step in order to ensure the lasting impact and interconnectedness of datasets would be to support further linked open data (for example, Prag and Chartrand 2019; Depauw 2019).

Several types of verbal multi-word expressions are comparative concepts (Haspelmath 2010) that are realised in language-specific ways. Consider for instance Coptic light-verb constructions that create statives (for example, ⲕⲁ-ⲣⲱ= *ka-rô*= "to keep silence" and ⲕⲁⲣⲁⲓⲧ *karaeit*) (Chapter 7), post-classical Greek ones that pragmaticalise into discourse markers (for example, ποίησον ἀγάπην *poiēson agapēn* "please") (Chapter 6), or verb-particle constructions and their relationship to Greek preverbation (for example, ἀπό-βαίνω *apo-bainō* "to go away") (Chapter 5). While the chapters of this volume are focussed on Indo-European and Afro-Asiatic languages, there are many corpus languages other than those treated in this volume which would proffer fertile ground for future work (cf. Fleischman 2000). We hope to have inspired those working on less well-resourced languages in general to consider multi-word expressions, such that comparative concepts can be refined and descriptive categories can be polished.

For Greek, Latin/Romance, and Egyptian, we have observed considerable variability, diversity, and development across time periods. In fact, verbal multi-word expressions, and the micro-variations in their lexicon and morphosyntax in particular, can often provide us with "sociocultural, economic, and geographical coordinates in time and space" (Tagliamonte 2025: 3), almost like a compass. This volume has focussed on the stages of data, methodology, and analysis specifically, yet we hope the compass we have created will now guide us towards the interpretation stage, that is, the domain of logical reasoning, the famous "why" question, and the critical assessment of multiple perspectives and avenues.

Acknowledgements

This work received support from the CA21167 COST action UniDive, funded by COST (European Cooperation in Science and Technology).

Bibliography

D'Angelo, Marzia, and Federica Nicolardi. 2024. Addressing Material Issues through Digital Solutions: Maque-IT and the Virtual Reconstruction of the Herculaneum Papyri. In Nicola Reggiani (ed.), *Digital Papyrology III: The Digital Critical Edition of Greek Papyri: Issues, Projects, and Perspectives*, 303–316. Berlin: De Gruyter. https://doi.org/10.1515/9783111070162-017

Depauw, Mark. 2019. Trismegistos: Optimisation Interoperability for Texts from the Ancient World. In Annamaria De Santis and Irene Rossi (eds.), *Crossing Experiences in Digital Epigraphy: From Practice to Discipline*, 193–201. Berlin: De Gruyter. https://doi.org/10.1515/9783110607208-016

Fendel, Victoria. 2024. Epilogue: Taking Wing. In Victoria Fendel (ed.), *Support-verb Constructions in the Corpora of Greek: Between Lexicon and Grammar?*, 327–340. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.14017941

Fleischman, Suzanne. 2000. Methodologies and Ideologies in Historical Linguistics: On Working with Older Languages. In Susan Herring, Pieter Reenen, and Lene Schøsler (eds.), *Textual Parameters in Older Languages*, 33–58. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.195.03fle

Haspelmath, Martin. 2010. Comparative Concepts and Descriptive Categories in Crosslinguistic Studies. *Language* 86(3). 663–687. https://doi.org/10.1353/lan.2010.0021

Hervais-Adelman, Alexis, and Laura Babcock. 2025. Magnetic Resonance Imaging (MRI). In Ana María Rojo López and Ricardo Muñoz Martín (eds.), *Research Methods in Cognitive Translation and Interpreting Studies*, 304–323. Amsterdam: John Benjamins. https://doi.org/10.1075/rmal.10

Leipoldt, Iohannes. 1913. *Sinuthii archimandritae vita et opera omnia iv*. CSCO 73/Scr. Copt. 5. Louvain: Durbecq. (reprinted 1954).

Marušić, Borislav. 2018. Besonderheiten der Funktionsverbgefüge in der deutschen Konzernsprache. *Jezikoslovlje* 19(1). 87–106.

Mcarthur, Tom, Jacqueline Lam-Mcarthur, and Lise Fontaine. 2018. *Oxford English Dictionary*. Oxford: Oxford University Press.

Prag, Jonathan, and James Chartrand. 2019. I. Sicily: Building a Digital Corpus of the Inscriptions of Ancient Sicily. In Annamaria de Santis and Irene Rossi (eds.), *Crossing Experiences in Digital Epigraphy: From Practice to Discipline*, 240–252. Berlin and Boston, MA: Mouton De Gruyter. https://doi.org/10.1515/9783110607208-020

Santiago, Julio, Purificación Meseguer, and Javier Valenzuela. 2025. Reaction Time. In Ana María Rojo López and Ricardo Muñoz Martín (eds.), *Research Methods in Cognitive Translation and Interpreting Studies*, 324–346. Amsterdam: John Benjamins. https://doi.org/10.1075/rmal.10

Tagliamonte, Sali. 2025. *Analysing Sociolinguistic Variation*. 2nd edn. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009403092

Index

abstract noun (AN) 211, 214, 218, 223–225, 232–233, 235, 238–239, 241–243, 252
  verbal 213–214, 238
aesthetic ideal, concept of 51, 62
αἵρεσις *haíresis* (“choice”) 18–19
Aktionsart 101, 111, 113, 212
alternation 102, 107, 122, 125, 139, 142, 166, 177, 179
ambiguous terminology 47–50, 55, 61–62
artificial intelligence (AI) 6, 18, 53
ataraxia (“freedom from disturbance”) 14, 22

base verb. *See* verb: base
body-part noun 76–77, 82, 85, 91
Bohairic 5, 159–160, 166–172, 178, 182

causativity 1, 6, 76, 78–79, 102, 113, 139, 141–143, 170, 182, 211–219, 222–225, 231, 233–238, 240, 242–244, 250, 254, 256, 258, 261–262, 265–266, 269–270, 276
causee and causer 102, 173, 213–214, 217, 223–224, 232, 235–238, 240–244
Christian context 132, 136, 141, 144, 154, 272, 283
classical period, concept of 51, 62
cognitive linguistics 251, 254, 278–279
compositionality 2, 23, 76, 94, 105, 109–110, 119, 122, 133, 163, 165–166
Coptic 2–3, 5, 73, 159–167, 169–173, 177, 179–183, 284–285
cranberry item 5, 56, 166

diachronic variation 2, 5–6, 16, 43, 53, 55, 59, 63, 94, 100, 126, 160, 170, 182, 207, 249, 272, 276
diatopic variation 16, 55
dolor 6, 237, 249, 251–262, 264–279. *See also* pain

Egyptian Pre-Coptic (EPC) treebank 73–74, 94
ἦν διδάσκων *ēn didaskōn* (“he was teaching”) 196–198, 201–204
Epicureanism 3–4, 13–16, 18–23, 25, 32–35
Epicurus 13–16, 20–23, 26–27, 32–34
event decomposition 19, 101
eventive noun 5, 171

FAIR data 42, 44, 63
Foucault, Michel 41, 50
Funktionsverbgefüge. *See* light-verb construction (LVC)

gatekeeping 50, 62–63
Gospel of John 189
Gospel of Luke 164, 179, 189–207, 215–217
Gospel of Mark 189–191, 193–197, 199–201, 204–207, 215–216, 243
group-internal terminology 41, 43, 47–50, 52, 55–56, 62–63

hapax legomenon 55–56, 190
Hebrew 3, 6, 180, 190, 196, 200, 211–214, 217–219, 222–225, 231–233, 236, 238, 240, 243–245
Herculaneum 3–4, 7, 13–16, 19–20, 33–35, 284

illocutive force 132, 151, 153, 155
imperative mood 93, 131, 135, 145–147, 149, 151–154, 177–178
imperfect tense 6, 189–207
indexicality 49, 59
intensity 249, 251, 255–256, 259, 263, 265–269, 271–272, 279
intersubjectivity 131, 145, 149, 154

καὶ ἐγένετο *kai egeneto* (“and it happened”) 198, 200–201
κενεαγγεία *keneaggeia* (“evacuant treatment”) 54–56, 63
χράομαι *khraomai* (“to use”) 57–61, 167

large language model (LLM). *See* artificial intelligence (AI)
Latin
  archaic 6, 249, 253–255, 257, 264, 271–272, 276–277, 279
  classical 3, 243–244, 249–250, 253–255,

257, 264, 271–272, 276–277, 279
late 217, 242, 244, 249–250, 253–254, 272, 275–279
lexical affix 160–164
Lexical Integrity Principle 160, 180, 183
lexicalisation 19, 74, 171, 177, 182, 252
light-verb construction (LVC) 3, 5, 75–76, 84, 159–160, 164–174, 176–180, 182, 211, 285

machine learning. *See* artificial intelligence (AI)
Maque-IT 18, 285
medieval Spanish 3, 6, 249–250, 253–254, 272, 275–279
metaphor 6, 47, 49, 60, 110, 173, 249–256, 263–272, 276–279
middle voice 56, 58–59, 138–139, 142
monoclausality 110
morphological number 133, 148–149, 154
multi-word expressions (MWEs)
adjectival and adverbial 85–89
functional 90–93
nominal 75, 81–85
verbal 1–2, 5–6, 54, 56, 62–63, 74–79, 88, 103, 105, 131, 159–160, 164–165, 167–172, 174–182, 284–285

narratology 3, 6, 190–191, 203–204, 206
New Testament (NT) 3, 6, 189–196, 199, 203–207, 213, 215–217, 234, 243, 283
nisba idiom 75, 82–83, 85–86, 88–92, 94
noun incorporation 160–164

object marking 174
Old Egyptian 3–4, 73, 76, 80, 82, 86, 88, 91, 94. *See also* Egyptian Pre-Coptic (EPC) treebank
Old Testament (OT) 6, 190, 200, 207, 213, 217, 224, 234, 238, 242–244

pain 4, 6, 22, 28, 226, 237, 249–279. *See also dolor*
papyrus 3–4, 7, 13–22, 27–28, 33–35, 54, 132, 136, 145, 284
PARSEME (PARSing and Multi-word Expressions) 4–5, 7, 19, 54, 56, 74, 76, 81, 86, 90, 94, 103, 105, 108, 126, 133, 165, 169
participle 6, 58, 84, 87–88, 138, 191–193, 196, 199, 203, 205, 207, 214, 233
particle 1, 3, 5, 44, 75–76, 93–94, 99, 101, 103–113, 115, 126–127, 149–151, 190, 197, 285
periphrasis 6, 160, 189–191, 193–196, 200–207, 212, 238
Philodemus of Gadara 15–16, 21, 23–34, 284
philosophy 13–14, 16, 22–23, 27, 33–34, 40, 55, 235, 244, 284
φυγή *phugē* (“flight”) 18–19
politeness 3, 131–132, 136, 151–153
polite question 135, 147, 153
polite request 135, 146, 153
polysemy 49, 133, 137, 148, 259
Polystratus 20–21, 32
πορεία *poreia* (“journey”) 55–61, 63
pragmaticalisation 132–133, 135–137, 145, 147–148, 151, 153–154
preverb 5, 100, 111–115, 117–122, 124–126
preverbation 99, 111–113, 118, 123, 285
Pyramid Texts 4, 73–94
Teti 3–4, 73–94
Unas’ 3–4, 73–94

Sahidic 159–160, 162–182
scorze (papyrology) 15, 17
search engine optimisation (SEO) 45, 47–48
Septuagint 3, 190, 196, 200, 211–212, 215, 217–218, 222–223, 225–226, 231–233, 236, 238–239, 244
Sketch Engine 3–4, 39, 42–44, 50, 54–56, 60–63, 285
source language 212, 214, 233
Spanish. *See* medieval Spanish
subject-specific terminology 41, 47–50, 55–56, 61–63
support-verb construction 1, 6, 54, 56, 58–61, 63, 132–134, 137–139, 141–143, 145–146, 148, 151–154, 211–214, 216–219, 222–231, 233–240, 242–244, 249–262, 264–273, 275–279
aspectual 254, 256, 260–261, 266–268
causative 212, 214, 216, 218, 236, 243–244, 254, 256, 261–262, 265–266
degree 254, 256, 262
syntactic integration 133, 135–136, 148, 151, 154

target language 214
telicity 105–108, 111–113, 122, 124–127, 160, 169, 179, 261, 267, 269
theta-role licensing 169–170, 172–173
transitivity 102, 104–105, 108, 110, 113, 121, 125, 133, 161, 163, 169, 175–177, 179, 182, 218, 231, 235, 240–242, 260, 266

univerbation 1, 5, 100–101, 103, 111, 126, 166, 170–172, 182

verb
  base 109, 211, 216, 241–242
  change-of-position 102, 104, 106–107, 110, 113, 115, 117–118, 120–122, 124–126
  change-of-state 101–105, 108, 110, 113–117, 119–122, 124–126
  manner-of-motion 106, 108, 113, 118
  of realisation 57–59, 139
  phrasal 1, 5, 99–100, 103, 111, 126
  pre-verbed 99–101, 109–112, 114–115, 117–119, 121, 123, 126. *See also* preverb
  separable 103
verb stem (Biblical Hebrew) 162, 223, 243
  Hifil 214, 217–220, 222–223, 226, 231, 233, 238, 240, 243
  Piel 219, 222, 230–231, 240, 243
verbal aspect 202
Villa dei papiri 14
visibility 3–4, 6, 41–43, 53–54, 56, 62–63, 283
Vulgate 3, 211–212, 214, 217–218, 222–223, 225–226, 231–233, 236, 238, 240, 242, 244

wordhood 5
word order 59, 100, 104, 146, 148, 151, 241–242

X-ray 18

About the Team

Alessandra Tosi was the managing editor for this book.

Adèle Kreager proof-read this manuscript and compiled the index.

Jeevanjot Kaur Nagpal designed the cover. The cover was produced in InDesign using the Fontin font.

Annie Hine typeset the book in InDesign. The main text and heading font is Noto Serif.

Jeremy Bowman produced the PDF, paperback, and hardback editions and created the EPUB.

The conversion to the HTML edition was performed with epublius, an open-source software which is freely available on our GitHub page at https://github.com/OpenBookPublishers

Lucy Barnes was in charge of marketing.

This book was peer-reviewed by a number of anonymous referees. Experts in their field, these readers give their time freely to help ensure the academic rigour of our books. We are grateful for their generous and invaluable contributions.

This book need not end here...

Share

All our books — including the one you have just read — are free to access online so that students, researchers and members of the public who can't afford a printed edition will have access to the same ideas. This title will be accessed online by hundreds of readers each month across the globe: why not share the link so that someone you know is one of them?

This book and additional content is available at https://doi.org/10.11647/OBP.0478

Donate

Open Book Publishers is an award-winning, scholar-led, not-for-profit press making knowledge freely available one book at a time. We don't charge authors to publish with us: instead, our work is supported by our library members and by donations from people who believe that research shouldn't be locked behind paywalls.

Join the effort to free knowledge by supporting us at
https://www.openbookpublishers.com/support-us

We invite you to connect with us on our socials!



BLUESKY
@openbookpublish
.bsky.social



MASTODON
@OpenBookPublish
@hcommons.social



LINKEDIN
open-book-
publishers



INSTAGRAM
openbookpublishers_

Read more at the Open Book Publishers Blog

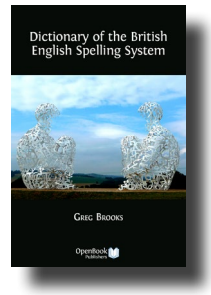
https://blogs.openbookpublishers.com

You may also be interested in:

Dictionary of the British English Spelling System

Greg Brooks

https://doi.org/10.11647/OBP.0053

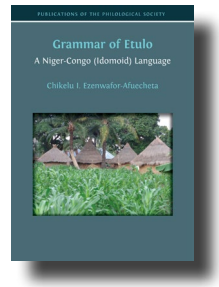


Grammar of Etulo

A Niger-Congo (Idomoid) Language

Chikelu I. Ezenwafor-Afuecheta

https://doi.org/10.11647/OBP.0467

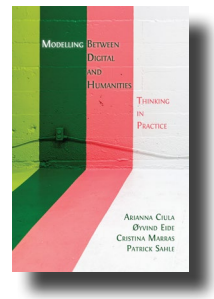


Modelling Between Digital and Humanities

Thinking in Practice

Arianna Ciula, Øyvind Eide, Cristina Marras and Patrick Sahle

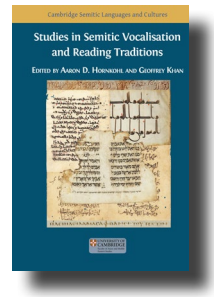
https://doi.org/10.11647/OBP.0369



Studies in Semitic Vocalisation and Reading Traditions

Aaron D. Hornkohl and Geoffrey Khan (eds)

https://doi.org/10.11647/OBP.0207



Verbal Multi-Word Expressions in Corpus Languages

Data, Methodology, and Analysis

Edited by Victoria Beatrix Fendel

This volume brings together a diverse, interdisciplinary cohort of scholars to shed new light on verbal multi-word expressions in corpus languages such as Greek, Latin, and Egyptian. Addressing in particular corpora related to in-groups, which are often overlooked due to entrenched research traditions or limited accessibility, the volume makes these corpora newly visible and methodologically relevant.

Contributions critically examine language-specific descriptive categories derived from comparative concepts used in cross-linguistic comparison and interrogate how corpus languages, with their non-expandable and historically bound data, challenge existing frameworks, language-internally, cross-linguistically, and typologically. Drawing on expertise from linguistics, computer science, papyrology, and Egyptology, *Verbal Multi-Word Expressions in Corpus Languages* explores both the theoretical and practical dimensions of studying verbal multi-word expressions, including cutting-edge digital solutions. Grounded in original data, the research pushes beyond current boundaries, advancing typological understanding and methodological precision.

The volume firmly supports multilingual research practices and challenges the dominance of English in academic discourse. The result is a critical, data-rich resource that advances our understanding of how verbal multi-word expressions function across historically significant but methodologically under-explored corpora.

Verbal Multi-Word Expressions in Corpus Languages will be essential reading for historical linguists focused on corpus methods and the syntax-lexicon interface; general linguists studying multi-word expressions comparatively; and digital humanities scholars developing tools for large-scale corpus analysis, as well as those interested in specific, underexplored corpora like those related to Galen and Hippocrates or the Herculaneum papyri.

This is the author-approved edition of this Open Access title. As with all Open Book publications, this entire book is available to download for free on the publisher's website. Printed and digital editions, together with supplementary digital material, can also be found at http://www.openbookpublishers.com.

Cover image: Photo by Matthew T. Ireland, all rights reserved
Cover design: Jeevanjot Kaur Nagpal

PhilSoc
The Philological Society

OpenBook Publishers

e book
ebook and OA editions also available
OPEN ACCESS