Erik Henriksson — Sonja Dahlgren — Marja Vierros

Phonological Variation in Greek Papyri

Two Case Studies Using PapyGreek Search

1 Introduction

Documentary papyri[1] have long been recognized as a remarkably direct source of insight into Postclassical Greek. From a linguistic perspective, these texts are invaluable for the various types of writing errors they contain, which can offer significant clues about the evolution of Greek during the Greco-Roman period.[2] Earlier, papyrologists had to rely on traditional reference texts such as those by Mayser and Schmoll and Gignac[3] to interpret these irregularities. However, digital advancements have reshaped the field,[4] and now that all published texts are digitized and accessible online (https://papyri.info), linguistic variants are accessible in a machine-readable format, enabling new methods of analysis.

In this study, we delve into the phonological aspects of linguistic variation in Greek documentary papyri using PapyGreek Search, a query tool developed as part of the PapyGreek project (https://papygreek.com/search).[5] Our focus is on the phonological variation that seems to be related to the prolonged contact between Greek and Egyptian-Coptic – a language contact situation that lasted for over a millennium after Alexander the Great's conquest of Egypt. Given that most papyrus findings originate from Egypt, the interaction between Egyptian (in its Demotic or Coptic stages) and Greek used in Egypt is relatively well-documented. Frequently, Greek scribes were Egyptians with varying levels of proficiency in Greek as a second language, leading to distinct errors in writing.[6] As Gignac first noted in 1976 and reaffirmed in 1991, many unconventional spellings in documentary Greek could be attributed to phonological transfer from Egyptian-Coptic due to bilingualism.[7] Utilizing our novel tool, we explore this interaction further, building on previous studies such as Dahlgren 2017 and 2020.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 758481).

1 Following common practice, the term 'papyri' in this study will also include ostraca and wooden tablets.
2 E.g. Dickey 2011.
3 Mayser – Schmoll 1970; Gignac 1976.
4 Reggiani 2017.
5 Henriksson – Vierros, this volume.
6 E.g. Vierros 2012.
7 Gignac 1976, 57–60; Gignac 1991.

https://doi.org/10.1515/9783111070162-012

The chapter proceeds as follows. Section 2 reviews the research tools available for linguists interested in studying phonological variation in Greek documentary papyri (2.1). We also discuss the main types of phonological variation found in the texts, with a focus on the particular contact features typical to Egyptian Greek, such as underdifferentiation of phonemes and coarticulation (2.2 and 2.3). Section 3 details our methodology, including data collection and the implementation of an algorithm for finding the character-level differences between linguistic variants. Our case studies are presented and evaluated in Section 4. In Section 5, we discuss our findings, offering preliminary conclusions and directions for future research.

2 Background

In this section, we first review the research tools available for linguists interested in studying phonological variation in Greek documentary papyri. We then present the primary phonological variations found in these materials, focusing on those possibly influenced by language contact. Finally, we discuss coarticulation, a recurring characteristic of the Egyptian variety of Greek.

2.1 Previous work and PapyGreek Search

Past research into the phonology of Greek documentary papyri is primarily found within two significant works: the phonology volume of Edwin Mayser's *Grammatik der Griechischen Papyri aus der Ptolemäerzeit* (1926–1938), updated by H. Schmoll,[8] and Francis T. Gignac's *Grammar of the Greek Papyri from the Roman and Byzantine Periods* volume 1.[9] These grammars offer extensive, manually compiled lists of orthographic variations in the texts, a practice also followed by Teodorsson 1977. While these resources remain valuable to papyrologists, they have notable limitations. First, their approach is mainly qualitative and selective, offering only broad statistical insights.[10] Furthermore, e.g. Gignac 1976 is overly cautious in labelling nonstandard features to result from language contact; Gignac 1991 states this more clearly but is lesser known as a source to explain the variation. Also, with the significant increase in published papyri since the 1970s, some of the earlier data and interpretations have inevitably become obsolete.

8 Mayser – Schmoll 1970.

9 Gignac 1976.

10 Gignac 1976, for instance, offers statistical information rather vaguely using vocabulary such as "sometimes", "occasionally", "frequently", "very frequently" throughout his work.

Turning to more recent work, the Trismegistos Text Irregularities database (TMTI)[11] emerged as the first online platform for examining linguistic variation within Greek documentary papyri. The search functionalities of TMTI mirror those of the PapyGreek Search's linguistic variation search features. Both tools, for instance, allow users to search for character-level editorial modifications, including deletions, additions, and replacements.[12] Furthermore, both allow users to tailor their searches based on context, such as the corrections of ε to η either before or after κ. However, the tools also differ in some important respects. Firstly, the TMTI platform, as detailed in its "Methodology" section online,[13] relies on data collected in 2014 and 2016, while PapyGreek Search takes advantage of the most recent editions, updated weekly from the papyri.info data repository (https://github.com/papyri). Moreover, PapyGreek Search provides a higher degree of flexibility than TMTI, which currently restricts users to a context of only one letter or a diphthong before and after the correction. In PapyGreek, the query context can be placed *anywhere* within the word, rather than strictly before or after the changed letters. This feature, along with several other options for refining the search, is made possible by the integration of Regular Expressions[14] into PapyGreek Search's parameters. Finally, PapyGreek Search features a timeline chart of the search results, facilitating the exploration of diachronic trends of the phenomenon under investigation (see Section 4 for examples). For a closer look at PapyGreek Search and its functionalities, see Section 3 below and Henriksson – Vierros, this volume.

2.2 Types of variation

Much of the variation found in Egyptian Greek is similar to the variation common in Koine Greek everywhere, related to internal Greek phonological processes. These processes led, for instance, to many vowel qualities fronting and/or raising to [i], which is the result in Modern Greek. However, Egyptian Greek had some peculiarities that were not found elsewhere. These include the underdifferentiation of the voiceless and voiced stops (/p, t, k/, /b, d, g/), which frequently resulted in having a voiceless variant where a voiced one should be, and vice versa, probably often due to hypercorrection. Such phenomena stem from the absence of the voiced stops /g, d/ in Coptic (Loprieno 1995; Layton 2000). Nevertheless, Egyptian scribes knew that they existed in the Greek they used for

11 See Depauw – Stolk 2015.

12 As Depauw – Stolk 2015, 212–6 point out, editorial practices have not been consistent in the long history of editing and publishing Greek papyri. Consequently, despite more recent efforts of standardization, some textual irregularities have not been encoded as such in the digitized texts, and are not detected by either TMTI or PapyGreek Search. An algorithm that would find these missing variants would be desirable.

13 https://www.trismegistos.org/textirregularities/methodology.php.

14 Aho 1991.

work, and occasionally added them where they did not belong because it seemed 'Greek'. Consider, for instance, the variation between /g/ and /k/ in Table 1. As can be seen from the examples, there are misspellings regarding this particular underdifferentiation of phonemes from the Ptolemaic period to quite late into the Roman period. In all, there are 566 instances in the PapyGreek database of γ /g/ being replaced with κ /k/, and 1.489 vice versa.[15] However, the many instances of standard /k/ having been replaced with /g/ are partly explained by the both Greek and Egyptian-Coptic ten-dencies of assimilating voiceless consonants to voiced ones, as the next word after εγ was often δίκης or δεξιῶν, and it is likely that the scribes learned the form as a semi-standard (Dahlgren 2017, p. 155). In any case, confusion between voiced and voiceless stops was not a regular feature of Greek language-internally, so the phenomenon in Egyptian Greek can be connected to the language contact situation.

Table 1: Examples of variation between /g/ and /k/.

Nonstandard	Standard	Document	Date	Provenance
κεωρκων	γεωργῶν	o.narm.5	150–225 AD	Arsinoites
κρυτωπωλων	γρυτοπωλῶν	bgu.1.9	276 AD	Arsinoites
θυκατηρ	θυγάτηρ	o.heid.332	101-200 AD	Thebes
μετηνεκκα	μετήνεγκα	upz.1.14	157 BC	Memphis
εκραφη	ἐγράφη	stud.pal.3.303	617 AD	Arsinoites
εγ	ἐκ	p.amh.2.46	113 BC	Pathyris

Greek used in Egypt also had other types of variation particular to the region. Some variations involved nonstandard markings of consonants, such as lambdacism, i.e., variation between the two liquid consonants /r/ and /l/, due again to (Egyptian-)Coptic not having /r/ at least in all dialects.[16] The quality of /r/ in Egyptian-Coptic is a topic we will explore in this chapter (Section 4.1). Frequently, however, the variation is concerned with vowel orthography, which seems to have had a mind of its own in Egyptian Greek and has been studied far and wide by both Greek scholars and Coptologists alike.[17] The vowel variation concerned practically all interchanges between e.g. /ai/ and /e, a/, /i/ and /e:, ei, y, oi/, /u/ and /y/, as well as between /o/ and /u/. This last case was again linked to the language contact situation but was related to the different stress patterns of Greek and Egyptian-Coptic, rather than phoneme inventories. We study this topic in Section 4.2.

15 The search parameters used to find variations between /k/ and /g/ are "form=-κ+γ" and "form=-γ+κ". The searches were conducted on 1st July 2023. (For an explanation of PapyGreek Search's query syntax, see Section 3.4 below.)

16 Peust 1999, 127–32.

17 Girgis 1966; Gignac 1976; Teodorsson 1977; Consani 1993; Torallas Tovar 2010; Horrocks 2010; Dahlgren 2017.

Much of this variation, naturally, is a product of the internal development of Greek. The phenomenon known as iotacism, or itacism, which resulted in the Greek front and close vowels merging into /i/ over time, certainly caused variation. This is evident not only in documents penned by native Greek (L1) writers but also those written by second language (L2) users. This would, of course, be related to such matters as bilingual language users or those with at least some competency in Greek becoming accustomed to hearing especially high-frequency words with the phoneme /i/, even when they had been previously pronounced with /e:/ or /ei/, as in the word ἐκεῖνος. This could result in nonstandard orthographic forms, such as those shown in Table 2, if the writer did not remember the orthographic standard.

Table 2: Example of variation related to iotacism.

Nonstandard	Standard	Document	Date	Provenance
εκινος	ἐκεῖνος	p.col.8.242	401–500 AD	Arsinoites

The papyrological documents, therefore, offer a rich body of evidence that can, using tools such as PapyGreek Search[18] or Trismegistos Text Irregularities,[19] be studied with the intent of pinpointing certain phases of the diachronic development of Greek. For instance, we can seek to establish the time when the variation between /i/ and /ei/ began to surge, suggesting a merger of these two phonemes, or when exactly the former Greek stops /d/ and /b/ started to develop into the voiced fricatives /ð/ and /β/ that we observe in Modern Greek.[20] Additionally, the search tools can be used to study the other language involved in the scenario, giving clues about its properties through the misspellings of the L2 Greek users. And, as much as there was variation seemingly related to iotacism, even that was often tied to the strongest element of Coptic phonology, consonant-to-vowel coarticulation.

2.3 Coarticulation

Coarticulation, a phenomenon studied in articulatory and acoustic phonetics, refers to the process of sounds adapting to the quality of the adjacent ones in continuous speech.[21] It can either be anticipatory or carryover, or both; this means that sounds can either affect the following sounds or that there can be phonetic residues of the previous sounds still remaining on the sounds that follow. For example, a vowel can be changed

18 Henriksson – Vierros, this volume.
19 Depauw – Stolk 2015.
20 Horrocks 2010, 112.
21 E.g. Hardcastle – Hewlett 2000.

in quality by the uvular consonant /q/, which is produced further back within the vocal tract, so that /i/ results in a retracted production similar to [e].

Coarticulation is part of all speech; it is impossible to speak without sounds overlapping one another to some extent. But coarticulation can be very language-specific, serving the distinctive needs of languages regarding their most important phonological contrasts.[22] As we know, in language contact, the L1 features of a language are often transferred onto the L2 used.[23] This includes L1 coarticulatory patterns. Frequently, when studying features of a contact language variety, the research ends up revealing as much or even more about the L1 features causing the variation than about any new patterns within the contact language itself. Dahlgren's 2020 study of Coptic vowel reduction, based on L2 Greek misspellings, serves as an example.[24]

In this study, we focus on coarticulation because it is a distinctly Egyptian-Coptic feature.[25] Coarticulation is not part of Greek diachronic development to the same extent in general, and more specifically, it is not similar in Greek as in (Egyptian-)Coptic, i.e. vowels do not adapt according to the consonantal context in the same phonetically systematic way as they do in (Egyptian-)Coptic. Egyptian-Coptic as a language belongs to the Afroasiatic language family, which has many other examples of the same type of consonant-to-vowel coarticulation in e.g. Arabic.[26] It is therefore a specific element that differentiates Egyptian Greek as a contact variety from the Greek-internal developments that formed the basis of the Koine Greek form used in all Greek-speaking areas from the Roman Period onward.[27]

3 Methodology

This section outlines the methodology used in our case studies examining phonological variation in documentary Greek. We begin by offering an overview of the data collection and preprocessing procedures utilized in PapyGreek Search. This is followed by a description of our method for identifying linguistic variants and the associated algorithms used for discerning character-level differences. Finally, we give a brief overview of our methods for storing and querying these variants.

22 Manuel 1999.
23 Weinreich 1968; Thomason – Kaufman 1988; Matras 2009 etc.
24 Dahlgren 2020.
25 Dahlgren 2017.
26 Bellem 2007; Ryding 2005.
27 Dahlgren 2022.

3.1 Data source and preparation

We use the texts from the Duke Databank of Documentary Papyri (DDbDP), a comprehensive digital repository containing more than 50,000 documentary papyri, accessible at https://github.com/papyri/idp.data. The encoding of these texts is based on the EpiDoc XML schema,[28] a system that utilizes various XML tags to manage the transcriptions and editorial handling of the texts. Our text preprocessing methodology is delineated by Henriksson – Vierros in this volume;[29] here it suffices to note that our linguistic variation database hinges on the modern editorial corrections embedded in the texts.[30] As an example, an irregular form like διδι ("give") might be regularized by an editor to διδε. These variant forms are contained by `<choice>` tags in the XML schema, with `<orig>` and `<reg>` elements nested within the parent tag representing the original and standardized words, respectively. What we are interested in are the character-level differences between the original and standardized forms (e.g., the "ι" in διδι corrected to "ε" in διδε), including the surrounding context of those differences. Therefore, we needed an algorithm capable of detecting these differences between the variant forms.

3.2 Identifying character-level variants

The search for a suitable algorithm for the task at hand required a comparison of various options. Our primary requirement was to find an algorithm that could find character-level differences which correspond to the editor's intended corrections, which is sometimes different from detecting *minimal* string differences from a computational perspective.[31] We also considered how different tools respond to ambiguous differences. For example, given the variants εγραφε and ενεγραφε, it appears evident to a human that the difference lies in the prefix εν-, which is absent from εγραφε. Yet, many tools would interpret this as the addition of an infix νε after the initial ε, which is clearly not the case here. After examining numerous algorithms, we found that Python's "difflib" library, utilizing gestalt pattern matching,[32] offered the most intuitive results. Therefore, we chose to use this library to find the character-level edits. For details of the process, see Section 3.3 in Henriksson – Vierros, this volume.

Our system treats the differences between the original and regularized forms as a series of *edit instructions*, essentially a guide on how to transform the original form into its corrected version. Four types of commands exist: copy, insert, delete, and replace. Take for instance the transition from δοραιαν to δορεαν. Our algorithm, built on the

28 Elliott – Bodard – Cayless *et a*l. 2006.
29 See also Vierros – Henriksson 2017 and Vierros 2018.
30 E.g. Stolk 2018.
31 For a detailed discussion, see Gusfield 1997.
32 Ratcliff – Obershelp 1988.

difflib-based string comparison method, suggests the following instructions: copy δορ, replace αι with ε, and then copy αν. We also consider the surrounding environment of the variants; in this example, αι is preceded by δορ and followed by αν, a potentially interesting context for this spelling error. Cases with multiple errors within a single word, like γιτοονος corrected to γειτονες, are managed by our system through a combination of copy, insert, delete, and replace commands.

3.3 Database structure

As explained above, our system handles the differences between the original and regularized forms through a series of edit instructions. For effective retrieval, these edits and their adjacent context are stored in an indexed MySQL database table, with dedicated text fields for each one ("original," "regularized," "original_before," "regularized_before," "original_after," "regularized_after"), along with a field indicating the edit operation in question ("copy," "insert," "delete," "replace"). Furthermore, to enable case and diacritic-insensitive searches, the table includes fields for de-accented and lowercase versions of these strings.

The variation table further includes a "token_id" field that links each indexed variation to its corresponding word in the database. This arrangement not only allows querying instances where an editor has altered a nonstandard form, but also enables more complex queries that integrate the change with morphology, syntax, and document metadata. This extended functionality of PapyGreek Search is elaborated in Henriksson – Vierros, this volume; here we primarily utilized its variation search function.

3.4 Search Queries for Variations

The PapyGreek Search interface is available at https://papygreek.com/search, where users are presented with a main search box where search terms can be entered using specific parameters. For the most basic word searches, the user would input "form=" followed by the desired term (without diacritics), such as "form=και" to find all instances of και. For queries specifically targeting linguistic variations, a special syntax is implemented. Users apply the "form" parameter in combination with symbols representing editorial actions: + (plus) for additions, - (minus) for deletions. As an illustration, "form=-ι" would yield all instances where ι has been deleted by an editor, while "form=+ι" would reveal all cases where ι was added.

Finding editorial replacements requires a combination of the - and + symbols. If one wishes to find instances where ε has been replaced by η, the search term would be "form=-ε+η". To further refine the search based on context, the symbols > (before) and < (after) are used. For instance, to find instances where ε has been corrected to αι following λ, the query would be "form=λ>-ε+αι". Finally, PapyGreek Search handles re-

gular expressions through the “regex:form=” parameter, which allows for the creation of more intricate and specific search criteria.

4 Case studies

In this section, we will illustrate how PapyGreek Search can yield novel and potentially significant results for previously unexplored phonological phenomena. There are two especially interesting cases that involve Egyptian-Coptic phonological impact visible in the nonstandard Greek orthography that show unclear or conflicting results: (a) how the adjacency of /r/ affects vowel quality, and (b) how the Egyptian-Coptic-influenced /o, u/ allophonic variation in Greek texts gets confused with nonstandard case inflection, and is usually taken to only mean morphological variation. We will start with the phenomenon of /r/ both fronting and retracting vowel quality, all the while still following a phonological symmetry in doing so: the actual nature of this consonant seems very liquid indeed in coarticulation, often merely transferring information of phonemes near it to other phonemes around it.

4.1 Case study (a): vowel variation related to the adjacency of /r/

The phoneme /R/, which in sociolinguistic (or sociophonetic) studies represents all variants associated with /r/-like sounds, regardless of what they are precisely phonetically, can be understood as a wider representative for all the phoneme qualities associated with what is written with one letter only in the world’s languages: <r>. This is a phoneme with a particularly wide range of variation for a consonant. It seems that /r/ can be almost anything that is an oral lingual sonorant consonant that is not particularly palatal (such as /j/), lateral (such as /l/), or bilabial (such as /m, b, p/). While there is phonetically no natural class that forms ‘rhotics’, they are nevertheless considered as part of the same phonemic group by most language speakers. What exactly is ‘rhoticity’ is unclear among phoneticians, as for instance a Spanish trill can be pronounced as a tap by some speakers, and these taps can be produced differently from one speaker to another; nevertheless, there is an intuitive connection between the various sounds written with the grapheme <r>, and the variants mentioned above are understood as allophones of the one and only /r/ in the language.[33]

Despite this one grapheme <r> being the only letter describing the phoneme in written language, /R/ actually has quite a few representatives phonetically, not all belonging to the same category. There is an alveolar trill /r/, as is in use in e.g. Finnish and Italian;

33 Scobbie 2006, 338–9; see also Rennicke 2015 for (socio)phonetically different variants of /R/ in Brazilian Portuguese

there is an alveolar tap (sometimes also called a flap) /ɾ/, which can be heard in e.g. American English when pronouncing a geminate /t/ as in *better* [bɛɾəɹ]; there is a retroflex approximant /ɻ/ as can be heard in e.g. Indian English (and generally in Indian native languages); and there is an alveolar approximant [ɹ] as can be heard in British English rhotic dialects, as well as in American English, when pronouncing e.g the word-final /r/ in *better*. In some languages, /r/ pairs up with /l/, both of which are called liquids, and these are not understood as individual phonemes; such is the case in e.g. Japanese. As can be seen from all this, /R/ is phonetically interesting.

In Greek misspellings coming from the Roman-Byzantine period Egypt, there is variation worth noting related to whether the adjacent vowel quality to /r/ is fronted or retracted, which both occur. Even more than in Greek texts, this is noticeable in Greek loanwords in Coptic texts, which naturally show even more integration to native language phonology than the second language (L2) Greek used by what presumably was mostly Egyptian writers. This is because loanwords are most often treated as native language words and for this reason, they have undergone adaptation to the native language phonological system.[34] But why does /r/ affect both directions?

Phonetically, /r/ belongs to the coronal consonant group, consonants formed with the tip (apical) or blade (laminal) of the tongue.[35] (European) Coronal consonants are listed in Table 3 below.[36]

Table 3: (European) coronal consonants.

z	s	ð	θ	ʒ	ʃ	n	d	t	ɹ	l	r	ɾ

Because coronal consonants are formed with the front part of the tongue at the front (from dental to postalveolar) part of the oral cavity, they usually raise or front the nearby vowel's quality; only in the case of retroflex phonemes, in which the tongue tip literally flexes backwards, vowel quality can actually be retracted as well.[37] On the other hand, in many languages of the world, also other types of /R/-sounds seem to retract vowel quality; this is the case, for example, in Modern Arabic[38] and Swedish.[39] In Swe-

34 E.g. Haugen 1950, 215–22; Weinreich 1968, 26–7; Major 2001, 136–7.

35 Ladefoged – Maddieson 1996, 11.

36 According to Dixon 2002, Australian aboriginal languages contain some coronals that are not found in consonant inventories of the European languages. These include for instance nasal and lateral consonants with a dental articulation, as they contrast with so-called peripheral (i.e. non-coronal; labial and velar) consonants, which is a typical feature of these languages. In this study we have only included coronal consonants that are relevant for the Greek study.

37 Flemming 2003, 335–6.

38 E.g. Ryding 2005, 26.

39 Riad 2014, 18–21.

dish, the quality of the liquid is the actual retroflex approximant /ɻ/, but in Modern Arabic, it is the same alveolar trill as in e.g. Finnish.

The answer to this seemingly curious feature of the phoneme use also in Egyptian-Coptic lies in the phonetic feature of /R/ universally. It is a somewhat 'weak' phoneme that eventually even tends to disappear from languages through lenition,[40] and this is probably also the reason it seems to affect vowel quality in two opposite directions. /R/ is (literally) so liquid in nature that it can take coarticulatory effect even from the syllables before or after the immediately adjacent phoneme, in anticipatory or carryover coarticulatory effect. In the case of consonant-to-vowel coarticulation, i.e. the process of vowels adapting to the manner or place of the adjacent consonants, this means that the adjacent vowel's quality can likewise be altered to follow the quality of the phoneme before or after /R/, the liquid merely transferring these phonetic properties onto the vowel. In effect, what this means is that if there is a front phoneme, consonant or vowel, before or after /R/, the vowel affected by /R/ in coarticulation will be fronted. If there is a back phoneme before or after /R/, the vowel next to it will be retracted in quality.

From what can be gathered from the evidence of this phenomenon in Egyptian Greek is that it seems that being the weak phoneme /R/ is, regardless of its precise quality in Egyptian-Coptic (or Greek), it is indeed itself affected by phonemes in the syllables around it, and thereafter transfers the phonetic properties of them onto the vowel qualities adjacent to it. The easiest phoneme variation pair with which to show this is eta/epsilon (in the Postclassical era phonemically ε, η /e, ę/[41]) due to this being among the only ones that cannot be confused with, for example, case variation, which is a very common problem when studying the round vowel (ο, ου /o, u/) variation. In addition, this goes toward proving that eta had not raised to /i/ at this point of the Egyptian Greek internal phonological development, as often assumed;[42] we find evidence both for synchronic contact effect and diachronic phonological development at the same time, with the same search parameters. In phonetic quality, eta represented the near-close front unrounded vowel /ę/, but its quality seems to have been positionally variable.[43] However, as there are, naturally, also other examples of the same fronting/retracting phenomenon connected to the adjacency of Egyptian Greek /r/, and instances where this happens near the other liquid consonant, /l/,[44] we have also included /i/ and /y/ into the search: in the examples, there are also instances in which even /i/ and /e/ are in variation, and /y/, naturally, is a frequent phonetic variant of /i/.

Some examples of the results of the search are presented in Table 4 below, and Figure 1 shows the timeline of the variation: it mostly takes place in the centuries of the emer-

40 See Rennicke 2015 for discussion and analysis.
41 Horrocks 2010, 112.
42 Teodorsson 1977; Horrocks 2010, 165–70.
43 Horrocks 2010, 167–70; Dahlgren 2017, 103–6.
44 Dahlgren 2017, 100–6.

gence and strongest use of Coptic.[45] This seems natural: at least some of the Egyptian L2 Greek writers used the same alphabetic writing system also in their own language, so they were familiar with the specific connection between phonology and orthography, creating misspellings that were based on forms originating from the spoken language.

In PapyGreek Search, the parameters used to examine variation adjacent to /r/ include using the regex mode and allowing /r/ to potentially occur with any vowel quality. However, we have intentionally omitted ο /o/, ω (phonemically /o:/ before the Postclassical era) and ου /u/. This decision was made to avoid the significant number of variants likely resulting from the merging of genitive and dative cases in the Postclassical era, which end in *-u* and *-o*, respectively. Since some of these seemingly case-related variants can also be phonetically-induced, such a search can be conducted separately to allow for a lighter phonemic analysis within those particular results. For the sake of brevity in this chapter, we have simply excluded this analysis, partly because there is never any certainty without looking into each instance separately to be able to decipher whether the variant could be seen as something phonetic (with similar misspellings occurring also elsewhere in the text) or whether it likely was regarding the famous case merger (with few phono-orthographic misspellings, but with other morphosyntactic variation present in the text). We have also excluded α /a/ from the search to avoid numerous examples of word-final vowel reduction to *schwa*, a phenomenon unrelated to the effect of the liquid consonant in the word, as in τεσσαρες from the standard τέσσαρας (e.g., bgu.4.1051, 30 BC–14 AD).[46] All in all, the search yields 887 entries in which the vowel quality is changed after the occurrence of /r/.[47] Run in the opposite direction, there are 1.460 variants of the vowel quality being changed before /r/. From this, we could deduce that Egyptian-Coptic coarticulation might have been more anticipatory in nature than carryover, but obviously reaching such a conclusion would require more extensive research on that precise phenomenon.

Table 4: Consonant-to-vowel coarticulation on the vowel following /r/.

Nonstandard	Standard	Document	Date	Provenance
κατακεχωρεκα	κατακεχώρηκα	cpr.1.198	138 AD	Arsinoites or Herakleopolites
μιλιαρισια	μιλιαρήσια	sb.6.9140	601 AD	Arsinoites (?)
βεριδαριου	βερεδαρίου	p.lond.4.1383	708-710 AD	Aphrodites Kome (Antaiopolites)
διατρεψαι	διατρῖψαι	bgu.4.1208	27-26 BC	Busiris (Herakleopolites)

45 Dahlgren 2017, 28–34.

46 See Dahlgren – Leiwo 2020 for this Egyptian-Coptic -induced phenomenon; also more in Case Study (b).

47 This search query targets vowel variations (ε,η,ι,υ) immediately following the consonant ρ, using the syntax "regex:form=ρ$>+^[εηιυ]$-^[εηιυ]$". Data was retrieved on July 1, 2023.

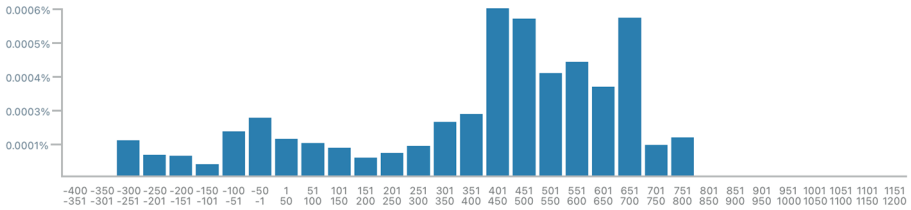

Fig. 1: Relative frequencies of words with vowel variation (ε, η, ι, υ) after ρ.[48]

In the first example shown in Table 4, κατακεχωρεκα from κατακεχώρηκα, the nonstandard production has an epsilon /e/ after /r/ instead of the standard eta /ę/. This is likely due to the retracting effect of /k/, a velar consonant, in the following syllable, which counts as one of the back consonants.[49] Μιλιαρισια from μιλιαρήσια, on the other hand, has a fronted nonstandard variant of the standard eta /ę/; it has raised to /i/. Given that the document dates to 601 AD, which is later than that of our first example, the change might be considered to reflect an advanced stage of completion of the raising of eta to /i/. However, there are many examples in Egyptian Greek documents of eta still being in variation with epsilon even in later centuries than this,[50] so it is equally likely that this is still phonetic variation. In general, the vowel quality in Greek used in Egypt was often retracted near /s/.[51] However, as the said /s/ is itself still followed by /i/ and the previous syllables contain two /i/'s, there is a very strong possibility of the combination having resulted in a raised vowel quality regarding eta, the liquid consonant just acting as a bridge between the front vowels.

Another, perhaps clearer and chronologically later example of a raised vowel quality adjacent to /r/ can be seen in the nonstandard variant βεριδαριου from the standard βερεδαρίου. Here, an /i/ replaces (epsilon) /e/. The vowel is positioned between /r/ and a front consonant, the coronal consonant /d/, which (together) evidently raise the vowel quality. Last, διατρεψαι from διατρῖψαι has a replacement of the standard word-medial iota with an epsilon, and in this example, the vowel quality seems to have been retracted by the combined forces of the /r/ before it and the bilabial (psi) after it. This earliest example, dating from the last pre-Christian century, should predate any poten-

48 Relative frequencies in the graph denote the proportion of word tokens from the search results in relation to the total number of word tokens in the entire papyri.info text database for each 50-year interval (e.g., 350–301 BC). Words from loosely dated documents spanning multiple 50-year periods were equally divided among these periods. For example, a document with 80 tokens covering a 200-year range allocates 20 tokens to each 50-year interval. For documents dated only before or after a specific time, the word count was distributed based on the average date range of 60 years, derived from fully dated documents in the database. Documents without any dating information were excluded from this calculation.

49 Jakobson 1968.

50 Dahlgren 2017.

51 Gignac 1976; see also Dahlgren 2017.

tial Coptic influence. In this particular instance, it is obviously difficult to say what role the liquid consonant may have had in the resulting variation as it could have been caused by the bilabial consonant alone (see more on the effect of the bilabials on vowels in the next section). Nevertheless, the examples above illuminate how varied results regarding a changed vowel quality can be found in the proximity of /r/.

4.2 Case study (b): vowel variation related to /o, u/ allophonic variation in Egyptian-Coptic

Greek documentary papyri have a substantial amount of variation between ο /o/ and ου /u/, which seems to be related to their allophonic status in (Egyptian-)Coptic.[52] Variation between these graphemes in Greek texts was already noted by Gignac,[53] who observed that this interchange often occurred in relation to stress: Greek unstressed ο /o/ was often replaced by ου /u/ and vice versa, which complies with the stress/allophonic rules of Coptic. Coptic had no unstressed ο /o/ (or /ɔ/) but it did have unstressed /u/,[54] and already Girgis remarked that Greek unstressed /o/ was often replaced with ου /u/ word-medially in Greek loanwords in Coptic.[55] In addition to this, this particular type of variation has been linked to the phonotactics of Coptic, i.e. the tendency of /o/ being replaced with /u/ in the adjacency of /m/ and /n/.[56] Although the coarticulation of consonants on vowels was a strong feature in Egyptian-Coptic, and the nasal/bilabial environments both have a tendency to raise the quality of open vowels crosslinguistically,[57] it is nevertheless very clear that this variation is not limited to these contexts. Native language prosody is typically one of the last elements to be lost by L2 speakers of a foreign language,[58] and it appears to be the case in L2 Greek writing in Egypt as well.

To confirm that this phenomenon is mainly related to stress and not coarticulation, there are some examples of variation that seem to indicate a change in the stress position. For instance, both λουγου /lugu/ in PSI VIII 884, 2 (390 AD) and κομιονται in BGU IV 1123, 6 (30 BC–14 AD) demonstrate variation related to the /o, u/ contact transfer. The standard forms of these words are λόγου /ˈlogu/ “word (gen.)” and κομιοῦνται “to take care of”, respectively.[59] As can be seen, in λουγου the first syllable’s stressed ο /o/ has been replaced with ου /u/, seemingly indicating that for the writer, this was the unstressed syllable; the genitive ending might have been learnt by heart due to its high

52 Dahlgren 2017, 83–4.
53 Gignac 1976, 211.
54 Peust 1999, 250–4.
55 Girgis 1966, 81–5.
56 Horrocks 2010, 112; Peust 1999, 238–40.
57 Beddor 1983, 2015; Flemming 2009, 82–4, 92.
58 Gut – Trouvain – Barry 2007; Matras 2009, 231–3.
59 See Dahlgren 2017, 153; in Postclassical Greek, the stress system had changed from having a primarily pitch accent to having dynamic word stress.

occurrence in Greek, being used in many patronymic forms. Similarly, in κομιονται the third syllable's original stressed ου /u/ has been replaced with ο /o/, as if to follow the Coptic phonemic distribution of /o/ being used as the rounded vowel in the stressed syllable, and /u/ in the unstressed one. This is probably due to transfer of Coptic stress rules, which fit in with the variant form's apparent word stress position.

Judging by the descriptions of other stress-timed languages, Coptic seems to have been one. It tended to place stress on one of the last two syllables of the word.[60] It also seems that, typically for stress-timed languages,[61] the stress was placed on the heavy syllable, at least in disyllabic words; perhaps more related to the word stress position, i.e. typically near the middle part of the word, in longer ones.[62] Variation, therefore, between ο, ου /o, u/ could be explained by (Egyptian-)Coptic stress rules and the phoneme distribution related to them. The vowel group α, ε, ο /a, e, o/ was also subject to neutralization in an unstressed position, especially word-finally.[63] /a, e, o/ variation concerned especially verb semantics, and could cause confusion over whether hybrid verb formations such as κερασεν (κέρασον) or πεμψεν (πέμψον) were to be interpreted as infinitives or imperatives.[64] Similarly, the stress position of the replaced vowel quality in κομιονται matches Coptic stress rules:

1. it is on one of the last two syllables;
2. it is in the middle part of the word, although four syllables can not be parted exactly in the middle;
3. it also happens to be the heavy syllable of the word with two consonants following the vowel.

On studying the phenomenon related to the /o, u/ variation in Egyptian Greek, we made four separate searches for the variation between ου/ω and ο/ου. The first two of these, standard ω replaced with ου, and vice versa, give many results that are usually interpreted to stem from the genitive and dative case merger, visible in the word-final variation of ου and ω (in Postclassical Greek, phonemically /u, o/); this was the position of Greek cases. The first search yields 1,348 results.[65] Searched the other way around there are 966 instances of standard ου replaced with nonstandard ω.[66] From the conso-

60 Loprieno 1995, 37; Peust 1999, 273.

61 Nübling – Schrambke 2004, 284–5.

62 Dahlgren 2017, 83–4, 153.

63 Dahlgren 2017, 62–6.

64 Dahlgren – Leiwo 2020.

65 The search targets variations where the original form has ου /u/ and the regularized form has ω /o/. The search query ("form=-ου+ω") was executed on 1st July 2023. Out of the 1,348 results, 1,161 are word final.

66 The data with the query ("form=-ω+ου") was retrieved on 1st July 2023. We note, however, that on some occasions the writers used (silent) *iota adscript* in connection with the ω, e.g. to mark the singular dative case (-ωι), and if we want the query to include these variants, we should use the regex query ("regex:form=-(ω|ωι)$+ου$"); this gives us 1,068 results. When we add a restriction of word final position ("regex:form=-(ω|ωι)$+ου$<^$"), we get 839 results, which means that only 229 instances are *not* word final.

nant environments, it is clear that some of the examples seem more like involving case merger, such as the nonstandard production δεισκου /deisku/.

In δεισκου from δείσκῳ, one of the many examples of what seem to be case confusion, the word-final stressed /o/ has been replaced with /u/, effectively changing a dative case to a genitive one. The nonstandard form also has a raised vowel quality from /o/ to /u/ after /k/, a velar consonant that, at least theoretically, should more retract the vowel quality, so there is no easy explanation to link it to coarticulation. The personal name misspelling Πετρου from what would have been the standard here, Πέτρῳ, on the other hand shows the possible effect of case merger, the position of stress and the related allophonic distribution of /o, u/ in Coptic, as well as consonant-to-vowel coarticulation: the standard /o/ could have been raised under the influence of the preceding /t/ and /r/, both coronal consonants. Πετρου is a prime example of how complicated it is to categorize variation in Egyptian Greek. Ορμου from ὅρμῳ looks like a clearer case of case merger as the preceding consonant is a bilabial /m/, although it primarily lowers high vowels instead of raising lower ones, which has happened here. Μαλλουṭον from μαλλωτὸν is interesting, especially from the point of view of the Greek stress position; below, we will talk about the possibility of transfer of stress, but in this example, the stress position seems to have been faithfully kept, and seems to be reflected in the replacement of the original ω /o/ to ου /u/, which is the unstressed rounded vowel in Coptic phonology. Consonant-to-vowel coarticulation is another possibility, of course, with the surrounding consonants /l/ and /t/ of the vowel being coronal ones. Also in αλλου /'allu/ from ἄλλῳ /'allo/, the word-final unstressed /o/ has been replaced with /u/ as per Coptic allophonic stress rules. In addition to the contact-induced stress connection, the coronal consonant /l/ could in this instance be raising the vowel quality. However, with the multitude of these types of cases altogether, stress-related variation seems a likely scenario for many, if not most, of the nonstandard vowel qualities because the word-final vowel, on which the case marking rests, is often unstressed in the Greek standard forms, especially in disyllabic words.

Some examples of the first search can be seen in Table 5, and again Figure 2 shows the distribution over the centuries – again highlighting a peak in the centuries when Coptic was used.

Table 5: Variation between ου and ω.

Nonstandard	Standard	Document	Date	Provenance
δεισκου	δείσκῳ	p.brem.24	116 AD	Hermopolis (?)
Πετρου	Πέτρῳ	p.brook.16	651-700 AD	Krokodilopolis (Arsinoites)
ορμου	ὅρμῳ	p.cair.isid.15	309-310 AD	Karanis (Arsinoites)
μαλλουṭον	μαλλωτὸν	p.cair.masp.1.67006v.	566-570 AD	Antinoopolis (?)
αλλου	ἄλλῳ	p.freib.2.8	144 AD	Unknown

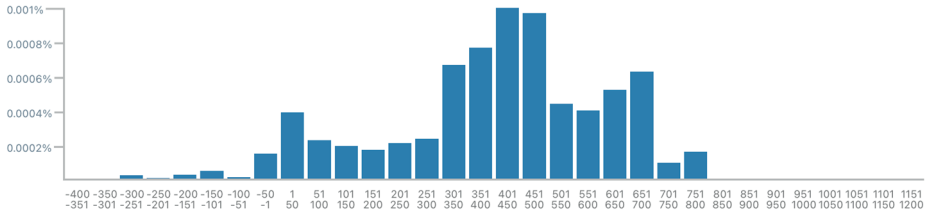

Fig. 2: Relative frequencies of instances of ου standardized to ω.

Table 6 shows the third search, related to the same variation of /u/ and /o/ but between the graphemes ου and ο, also proving that the quantity difference distinguishing *omicron* and *omega* had disappeared. Interestingly, the timeline in Figure 3 shows a dif-ferent distribution than in Figure 2 related to omega and ου: in Fig. 2, the variation is predominantly visible in the 3rd to 4th centuries AD, giving firmer evidence of the group belonging to the actual case merger category. In Fig. 3, there is variation both in the pre-Christian centuries as well as after; a high peak in the first century AD and a steady peak in the 3rd to 6th centuries AD. All examples show a possibility of stress transfer involvement, as well as a possible coarticulatory effect. The search yields 578 tokens.[67]

Table 6: Variation between ου and ο.

Nonstandard	Standard	Document	Date	Provenance
τουτου	τοῦτο	p.grenf.2.30	102 BC	Pathyris
εικουσι	εἴκοσι	p.ant.1.42	557 AD	Lenaiu (Antinoites)
μενουντος	μένοντός	basp.51.49	345 AD	Oxyrhynchos

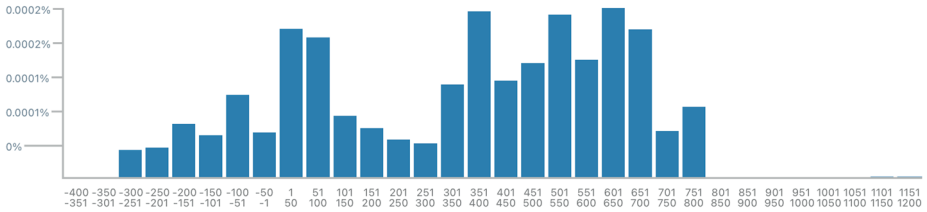

Fig. 3: Relative frequencies of instances of ου standardized to ο.

67 This search aims to find variants between ου /u/ and ο /o/ where the editor has corrected the nonstandard ου to ο. At the character-level, this corresponds to the editorial deletion of υ after ο, which in PapyGreek Search can be expressed as "regex:form=ο$>-^υ$". The search was conducted on 1st July 2023.

The fourth search, concerning the variation between omicron and ου but in another direction i.e. the standard ου being replaced with nonstandard ο, gives results that could, again, result from a number of things from case merger to coarticulation, but do also show a replacement of vowels that match Coptic stress rules. In μερος from μέρους, the replaced ο from the standard ου is on the second syllable, which is heavier, thus matching the stress position of the word that would have been more natural for Coptic, as do the next examples, αποδοναι, ησυχοντος, and ετομεν.[68] Figure 4 gives the distribution of tokens again, showing a peak in the later centuries from 4th to 6th, when Coptic was in use. The search gives 867 tokens.[69]

Table 7: Variation between *omicron* and ου.

Nonstandard	Standard	Document	Date	Provenance
μερος	μέρους	bgu.1.251	81 AD	Soknopaiu Nesos (Arsinoites)
αποδοναι	ἀποδοῦναι	bgu.2.595	75-85 AD	Arsinoites
ησυχοντος	ἡσυχοῦντος	sb.6.9138	576-600 AD	Arsinoites
ετομεν	αἰτοῦμεν	sb.6.9194	276-300 AD	Alexandria

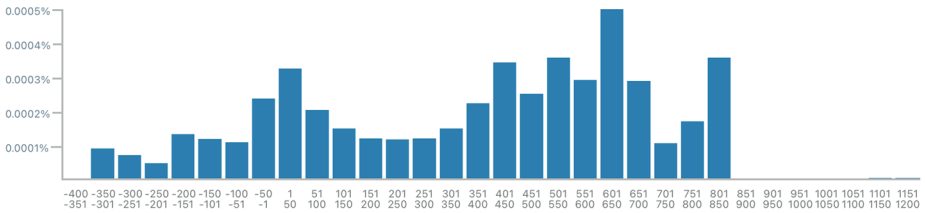

Fig. 4: Relative frequencies of instances of ο standardized to ου.

5 Discussion

In this chapter, we utilized the novel PapyGreek Search tool to examine phonological variation within Greek documentary papyri, focusing specifically on variations associated with the language contact situation between Greek and Egyptian-Coptic. We presented two case studies that explore potentially complex phonological variations. The

68 Dahlgren 2017, 133–8.

69 This search targets instances of nonstandard ο /o/ in place of the standard ου /u/; in other words, editorial additions of υ after ο. The search ("regex:form=ο$>+^υ$") was executed on 1st July 2023.

types of variation we chose for the study are notably challenging to interpret from textual evidence, in comparison to, for example, exploring more straightforward phonetic coarticulation that can be directly seen in the misspellings.

The first case study examined the variation of vowel qualities adjacent to /r/ in Greek documentary papyri. Through the analysis of specific examples, the study provided evidence suggesting that liquids may function as phonetic bridges in Egyptian Greek, transferring the phonetic properties of neighboring phonemes to the vowel qualities near /r/. The observations discussed in this study were obtained using the regex feature of PapyGreek Search. This feature allowed for the targeting of multiple vowel variations using a single search query, a task that was not achievable using traditional methods or previous digital tools. However, it was also acknowledged that determining the exact causes of these variations is often challenging, as writing errors can stem from various factors including, but not limited to, coarticulation, language contact, or phonological changes. Moreover, we focused on individual examples, and future research would benefit from a larger sample size and statistical analysis to enhance the credibility of the analysis. Nonetheless, the study underscored the capabilities of PapyGreek Search, paving the way for further investigation into the phonological characteristics of Egyptian Greek, and through contact-induced transfer effects, also Coptic phonology.

Second, we scrutinized vowel variation in cases, a subject that still remains inconclusive and multifunctional.[70] Still, phonetic and phonological factors must not be discounted when they coexist with morphological variation. Notably, the variation between /o/ and /u/ was relatively rare before the Roman period — precisely when Coptic began to make an appearance. Few instances were recorded during the Ptolemaic period, with the real influx of this variation starting in the Roman period, evident in Greek texts and in Coptic renderings of Greek loanwords.[71] This suggests a potential link to (Egyptian-)Coptic phonological influence in some sort of a bilingual milieu, whether this be related to the spoken level or orthographic practices. For L2 Greek speakers, the distinction between vowel qualities may not have been audibly discernible, but there might be a learned practice on the level of orthography to use omicron only for the stressed rounded vowel quality. This, however, is something we will not be able to completely verify within text linguistics because it rests on evidence from actual spoken language, which remains beyond our reach. Nevertheless, we believe our case study illustrates that certain traces of the spoken language can be inferred, even when dealing with such complex phenomena as stress transfer. With the aid of a more extensive sample size, diverse search parameters, and statistical analysis facilitated by PapyGreek Search, Coptic stress patterns could possibly be exhaustively extracted from the L2 Greek data from Egypt.

70 See e.g. Stolk 2015.

71 Gignac 1976, 207 n. 2.

The main value of PapyGreek Search, for the purposes of the present study, was to serve as a fresh interface to the already existing data found in XML-encoded source files, specifically pairs of irregular word forms and their editorial corrections. The construction of this novel dataset involved using an algorithm designed to discern the character-level differences between the original and regularized word forms, with the identified "edit instructions" being stored in a MySQL database for efficient retrieval. This is not the first interface designed for the purpose of finding text irregularities in documentary Greek.[72] However, we believe that having several similar tools is beneficial to the field of digital papyrology. Used in conjunction, they can either validate results or raise questions on the findings, encouraging careful scrutiny and planning of data collection and search parameters.

Looking ahead, the variation search functionality in PapyGreek Search could be advanced in several ways. Firstly, our database currently includes only documentary papyri, representing just one of the openly available collections where linguistic variation is encoded in a machine-readable format. By expanding the PapyGreek Database to include other collections encoded in EpiDoc XML – such as literary papyri and many epigraphic documents – the search tool could offer a broader perspective on linguistic variation across different text types and linguistic registers. Secondly, it would be highly valuable if the tool could detect not only those irregularities that have been corrected and encoded in the texts but also the numerous non-standard forms that have gone unnoticed by editors.[73] One approach to achieve this could involve training an ancient Greek language model to identify and correct nonstandard spellings and grammatical forms. This could uncover a vast number of previously unknown textual irregularities and potentially significantly enhance our understanding of how Greek evolved during the Greco-Roman period.

Bibliography

Aho, A. V. (1991), *Algorithms for Finding Patterns in Strings*, in *Handbook of Theoretical Computer Science, Volume A: Algorithms and Complexity*, ed. by J. Van Leeuwen, Amsterdam, 257–300.

Beddor, P. S. (1983), *Phonological and Phonetic Effects of Nasalization on Vowel Height*, PhD Diss., Bloomington (IL).

Beddor, P. S. (2015), *The Relation Between Language Users' Perception and Production Repertoires*, in *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS1041.pdf.

Bellem, A. (2007), *Towards a Comparative Typology of Emphatics*, PhD Diss., London.

Consani, C. (1993). *La koiné et les dialectes grecs dans la documentation linguistique et la réflexion métalinguistique des premiers siècles de notre ère*, in *La koiné grecque antique. I Une langue introuvable?*, Nancy, 23–39.

72 Cf. Depauw – Stolk 2015.

73 Stolk 2018.

Dahlgren, S. (2017), *Outcome of Long-Term Language Contact: Transfer of Egyptian Phonological Features onto Greek in Graeco-Roman Egypt*, PhD Diss., Helsinki.
Dahlgren, S. (2020), *The System of Coptic Vowel Reduction: Evidence from L2 Greek Usage*, Italian Journal of Linguistics 32(1), 211–27.
Dahlgren, S. (2022), *Egyptian Greek: A Contact Variety*, in *Ancient Indo-European Languages between Linguistics and Philology*, ed. by M. Bianconi – M. Capano – D. Romagno – F. Rovai, Leiden, 115–52.
Dahlgren, S. – Leiwo, M. (2020), *Confusion of Mood or Phoneme? The Impact of L1 Phonology on Verb Semantics*, in *Postclassical Greek: Contemporary Approaches to Philology and Linguistics*, ed. by D. Rafiyenko – I. Serẑant, Berlin – Boston, 283–301.
Depauw, M. – Stolk, J. (2015), *Linguistic Variation in Greek Papyri: Towards a New Tool for Quantitative Study*, GRBS 55, 196–220.
Dickey, E. (2011), *The Greek and Latin Languages in the Papyri*, in *The Oxford Handbook of Papyrology*, ed. by R. Bagnall, Oxford, 149–69.
Dixon, R. M. W. (2002). *Australian Languages: Their Nature and Development*, Cambridge.
Elliott, T. – Bodard G. – Cayless, H. *et al*. (2006). *EpiDoc: Epigraphic Documents in TEI XML*, https://epidoc.stoa.org.
Flemming, E. (2003), *The Relationship Between Coronal Place and Vowel Backness*, Phonology 20(3), 335–73.
Gignac, F. T. (1976), *A Grammar of the Greek Papyri of the Roman and Byzantine Periods*, I (*Phonology*), Milan.
Gignac, F. T. (1991), *Phonology of the Greek of Egypt, Influence of Coptic On the*, in *The Coptic Encyclopedia*, ed. by A. Suryal Atiya, New York, VIII, 186–8.
Girgis, W. A. (1966), *Greek Loan Words in Coptic (Part II)*, Bulletin de la Société d'Archéologie Copte 18, 71–96.
Gusfield, D. (1997), *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*, New York.
Gut, U. – Trouvain, J. – Barry, W. J. (2007), *Bridging Research on Phonetic Descriptions with Knowledge from Teaching Practice: The Case of Prosody in Non-Native Speech*, in *Non-Native Prosody. Phonetic Description and Teaching Practice*, ed. by J. Trouvain – U. Gut, Berlin, 3–21.
Hardcastle, W. J. – Hewlett, N., (2000), eds., *Coarticulation: Theory, Data, and Techniques*, Cambridge.
Haugen, E. (1950), *The Analysis of Linguistic Borrowing*, Language 26(2), 210–31.
Horrocks, G. (2010), *Greek: A History of the Language and Its Speakers*, 2nd ed., Malden (MA).
Jakobson, R. (1968), *Child Language, Aphasia and Phonological Universals*, The Hague [1941].
Ladefoged, P. – Maddieson, I. (1996), *The Sounds of the World's Languages*, Oxford.
Layton, B. (2000), *A Coptic Grammar*, Wiesbaden.
Loprieno, A. (1995), *Ancient Egyptian. A Linguistic Introduction*, Cambridge.
Major, R. C. (2001), *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*, Mahwah.
Manuel, S. (1999), *Cross-Language Studies: Relating Language-Particular Coarticulation Patterns to Other Language-Particular Facts*, in *Coarticulation. Theory, Data and Techniques*, ed. by W. J. Hardcastle – N. Hewlett, Cambridge, 179–98.
Matras, Y. (2009), *Language Contact*, Cambridge.
Mayser, E. – Schmoll, H. (1970), *Grammatik Der Griechischen Papyri Aus Der Ptolemäerzeit. Laut- Und Wortlehre: Einleitung Und Lautlehre*, Berlin.
Nübling, D. – Schrambke, R. (2004), *Silben- versus akzentsprachliche Züge in germanischen Sprachen und im Alemannischen*, in *Alemannisch im Sprachvergleich. Beiträge zur 14. Arbeitstagung für alemannische Dialektologie in Männedorf (Zürich) vom 16. – 18.9.2002*, ed. by E. Glaser – P. Ott – R. Schwarzenbach, Stuttgart, 281–320.
Peust, C. (1999), *Egyptian Phonology: An Introduction to the Phonology of a Dead Language*, Göttingen.
Ratcliff, J. W. – Metzener, D (1988), *Pattern Matching: The Gestalt Approach*, Dr. Dobb's Journal 46.
Reggiani, N. (2017), *Digital Papyrology I. Methods, Tools and Trends*, Berlin – Boston.
Rennicke, I. (2015), *Variation and Change in the Rhotics of Brazilian Portuguese*, PhD Diss., Helsinki.
Riad, T. (2014), *The Phonology of Swedish*, Oxford.

Ryding, K. C. (2005). *A Reference Grammar of Modern Standard Arabic*, Cambridge.
Scobbie, J. (2006), *(R) as a Variable*, in *Encyclopedia of Language and Linguistics*, X, 337–44.
Stolk, J. (2015), *Case Variation in Greek Papyri: Retracing Dative Case Syncretism in the Language of the Greek Documentary Papyri and Ostraca from Egypt (300 BCE–800 CE)*, PhD Diss., Oslo.
Stolk, J. (2018), *Encoding Linguistic Variation in Greek Documentary Papyri: The Past, Present and Future of Editorial Regularization*, in *Digital Papyrology II. Case Studies on the Digital Edition of Ancient Greek Papyri*, ed. by N. Reggiani, Berlin – Boston, 119–37.
Teodorsson, S.-T. (1977), *The Phonology of Ptolemaic Koine*, Gothenburg.
Thomason, S. G. – Kaufman, T. (1988), *Language Contact, Creolization, and Genetic Linguistics*, Berkeley (CA).
Torallas Tovar, S. (2010), *Greek in Egypt*, in *A Companion to the Ancient Greek Language*, ed. by E. Bakker, Malden (MA) – Oxford – Chichester, 253–66.
Vierros, M. (2012), *Bilingual Notaries in Hellenistic Egypt. A Study of Greek as a Second Language*, Brussels.
Vierros, M. (2018), *Linguistic Annotation of the Digital Papyrological Corpus: Sematia*, in *Digital Papyrology II. Case Studies on the Digital Edition of Ancient Greek Papyri*, ed. by N. Reggiani, Berlin – Boston, 105–18.
Vierros, M. – Henriksson, E. (2017), *Preprocessing Greek Papyri for Linguistic Annotation*, in *Journal of Data Mining and Digital Humanities. Special Issue on Computer-Aided Processing of Intertextuality in Ancient Languages*, ed. by M. Büchler – L. Mellerin, http://jdmdh.episciences.org/paper/view/id/1385.
Weinreich, U. (1968), *Languages in Contact: Findings and Problems*, The Hague [1953].