



Universiteit
Leiden
The Netherlands

Orthographic variation and register in the corpus of Greek documentary papyri (300 BCE-800 CE)

Stolk, J.V.; Bentein, K.; Janse, M.

Citation

Stolk, J. V. (2020). Orthographic variation and register in the corpus of Greek documentary papyri (300 BCE-800 CE). In K. Bentein & M. Janse (Eds.), *Varieties of post-classical and Byzantine Greek* (Vol. 331, pp. 299-326). De Gruyter Mouton.
doi:10.1515/9783110614404-012

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3512117>

Note: To cite this publication please use the final published version (if applicable).

Joanne Vera Stolk

12 Orthographic variation and register in the corpus of Greek documentary papyri (300 BCE–800 CE)

Abstract: The corpus of Greek documentary papyri from Egypt consists of various types of documents, such as letters, contracts and accounts, showing different types of linguistic variation. The concept of register is applied here to examine the relationship between the presence of non-standard orthography and the situational context according to the situational variables setting, participants, genre and production circumstances. Quantitative study shows that the participants involved and the genre of the document are predictors for the amount of orthographic variation that is found in a document. Qualitative analysis of the documents in a number of archives reveals that there are also other important factors, such as the choice of scribe, method of production and the stage of composition of the text that is preserved to us, to explain the presence of orthographic variation in the corpus of documentary papyri.

1 Introduction

The corpus of Greek documentary papyri from Egypt is known as rich source for linguistic variation. This corpus of more than 50.000 texts is available online via the PN and includes documents of various types, such as letters, petitions, contracts, accounts and lists, mostly written in Egypt between 300 BCE and 800 CE. The attested variation in spelling has been used to reconstruct the pronunciation of the spoken language at the time,¹ but the distribution of orthographic variation within the corpus may also be governed by language external factors.² Documentary papyri do not form a homogeneous corpus: different types of variation at various levels of the language are found across variety of document types. The concept of register could be applied to explain the distribution of variation in more detail, but the variables governing the variation at the level of orthography and morphology may not be identical to those at the level of syntax and pragmatics (§1.1). One might expect that standard orthography is a feature of a higher register, but non-standard orthography does

1 E.g. Gignac (1976), Teodorsson (1977).

2 Cf. e.g. Rutkowska & Rössler (2012).

not exclusively appear in informal contexts in the papyrological corpus (§1.2). In this paper, I will test several situational variables to see to what extent they govern the distribution of orthographic variation in the corpus of documentary papyri. After a short introduction to the research question, the corpus and methodology will be explained in more detail in Section 2. The quantitative analysis of standard and non-standard orthography in Section 3 is based on the editorial regularizations provided in digital editions (TMTI) and followed by a qualitative analysis of the distribution of orthographic variation in several archives in Section 4. A conclusion and discussion of the results is provided in Section 5.

1.1 Register and linguistic variation

Register variation is often understood in the form of a continuum.³ In some studies the whole corpus of documentary papyri is situated at the lower end of this register continuum,⁴ whereas other studies have demonstrated that a register continuum can also be found within the corpus itself and even within sub-corpora such as papyrus archives. For example, Bentein (2015a: 479) concludes about the linguistic variation in the so-called archive of the “Katochoi of the Sarapieion” from the second century BCE (TM Archive ID 119): “In terms of the earlier-mentioned register-continuum, we can say that the language of the dreams is situated most to the left (‘the low register’), followed by the letters and petitions respectively. This register-continuum itself can be seen as the sum of a number of linguistic dimensions (phonology, semantics, morphology, syntax).”

The concept of register to explain variation within the papyrological corpus has mainly been shown useful in the domains of syntax and pragmatics, such as for variation between complementation patterns⁵ and the use of particles.⁶ Comparison of features at other linguistic levels, however, does not always give the same results in individual texts.⁷ This is also a concern expressed by Halla-aho (2010: 172) about the Latin language in the papyri: “There is often an implicit assumption that a given letter would, as a whole, belong to a certain linguistic variety [. . .] I shall argue that in a given letter, different levels of language

³ Cf. Biber (1995: 31).

⁴ Bentein (2012, 2013).

⁵ James (2007), Bentein (2015c, 2017).

⁶ Clarysse (2010), Bentein (2015b).

⁷ Bentein (2015a: 479).

organization (phonological/orthographic, morphological, and syntactic) need not, and often do not, consistently relate to one linguistic variety (register or soci-lect).” This inconsistency may be caused by a difference between the external factors governing the spread of syntactic and pragmatic features and the ones governing orthographic and morphological variation.

1.2 Language and context

Linguistic register is the result of an interplay between “linguistic behaviour” and the “sociolinguistic context”.⁸ James (2007: 35–36) defines three levels of language in the papyri, namely “high (official), middle (some official, business, and some personal), and low (personal texts which seem substandard even within the Koine)”. The sociolinguistic context is here reduced to a cline of subject matters: from official and business related to personal texts. Linguistic behaviour is described in terms of its standardness or rather substandardness within the Koine. Characterisation of language in terms of vernacularity is not uncommon in historical linguistics, compare the following characterisation of the “formal” register by Hickey (2010: 8): “The lower this [vernacularity, JVS] is the more formal the register and hence the more standard the language will be [. . .] A high level of vernacularity implies a high incidence of non-standard features which are indicated by unexpected spelling and grammar.”

More formal or “high” register texts are thus expected to show more standard language features, while non-standard spelling and grammar is taken as a sign of a “low” register text. These divisions across the register continuum are visualized in Table 1.

Table 1: Context-based and language-based approach to a register continuum.

Continuum	“high”	“middle”	“low”
Context-based	official texts	business texts	personal texts
Language-based	low number of non-standard features	average number of non-standard features	high number of non-standard features

⁸ Porter & O’Donnell (2010: 297).

One would assume that the results of the context-based approach largely overlap with the language-based approach, but this expectation is not always born out in the case of papyrus documents. James (2007: 36) mentions several problematic examples: “Some early texts, which are official in content, such as the copy of the letter of Claudius to the Alexandrians preserved in P.Lond. VI 1912 (= Sel.Pap. II 212), can show obviously low-level or ‘late’ orthography and morphosyntactic features.”

Thus, individual examples show that non-standard language features can also be found in official contexts. In this paper I will test this relation between linguistic behaviour and sociolinguistic context in the whole corpus of published documentary papyri, focussing on the use of orthography. Are the documents that we would judge as formal or official in a context-based approach indeed the ones written with the most standardized orthography?

2 Approach to the corpus

Orthographic variation has not been annotated as such in digitalized papyrus editions, but the corpus offers an approximate measure of orthographic variation in the form of editorial regularizations (§2.1). This large scale identification of non-standard orthographic features allows us to do a quantitative analysis of the relation between orthography and the situational context. Different situational variables are selected (§2.2) and annotated in the corpus (§2.3) for this quantitative analysis (results in §3), whereas some other possibly relevant variables will be tested qualitatively in different sub-corpora in Section 4.

2.1 Counting orthographic variation

Digital editions of all published papyri are available in the PN, including regularizations of orthography and morphology in the text and apparatus. Together with Mark Depauw (KU Leuven), I collected and annotated more than 130.000 attestations of editorial regularization in a total corpus of almost 52.000 documentary papyri.⁹ The results can be accessed through the TMTI. All regularizations have been categorized into the type of regularization (such as o instead of ω, omission of ς, addition of ν) and annotated for linguistic level. For this study,

⁹ See Depauw & Stolk (2015). All results are based on the state of the PN of January 2014 and the state of TM in November 2017.

I only used regularizations involving an interchange, omission or addition of a single character or diphthong. The majority of these are the result of phonological changes in Post-classical Greek.¹⁰

Standard linguistic features can be identified as “linguistic elements which have a regular distribution”, but they can also be based on the “distribution according to occasion of use” (Milroy & Milroy 1985: 117). In orthography, the first type of standard features would be the form that is regularly used in contemporary sources and the second type is the form which is used in higher registers (e.g. more formal or literary language). Both types of definitions are used to regularize spelling in papyrus documents.¹¹ Editorial regularization has been applied to papyrus editions from the beginning of the twentieth century, initially meant to make papyrus editions easier to understand for readers used to Classical Greek. Editorial regularizations, therefore, tend to follow the norms of Classical Greek orthography (high register) and/or contemporary parallels (regular distribution). This means that they can give us an approximate measure for the amount of standard and non-standard orthographic features in a text. Of course, this method cannot have been entirely consistent throughout a century of scholarship and critical notes can be made to the linguistic accuracy of this old-fashioned system of annotation.¹² On the other hand, most of the volumes of papyrus editions contain a variety of documents and there seems no reason to think that the principles of regularization will differ significantly with respect to the original situational context of the document at hand. Qualitative analysis of a sample of texts from relevant archives (§4) will allow me to reassess and refine the quantitative conclusions based on these editorial regularizations (§3).

2.2 Selecting situational variables

The term register has been employed by Halliday, McIntosh and Stevens (1964: 87) in order to distinguish language varieties according to *use*, i.e. situationally defined varieties, from varieties according to *user*, such as dialects or sociolects. In this paper, I will follow the definition by Biber and Conrad (2009: 6): “a *register* is a variety associated with a particular situation of use (including particular communicative purposes). The description of a register covers three major

¹⁰ See Gignac (1976).

¹¹ Stolk (2018: 130–131).

¹² See Stolk (2018).

components: the situational context, the linguistic features, and the functional relationships between the first two components.”

There are multiple ways to define the situational context and the term “register” can be applied with different levels of generality, although the basic features of the communicative setting and the roles of the participants seem to take an important place in many approaches.¹³ Biber and Conrad (2009: 39–47) propose to describe the situational context according to the following parameters:

- (1) participants, including the number of participants and their social characteristics
- (2) the nature of the relationships among the participants
- (3) channel: mode and medium of communication
- (4) production and comprehension circumstances
- (5) setting: time and place of communication and historical time period
- (6) communicative purposes
- (7) topic or subject matter

Not all of these characteristics are equally relevant in every corpus study. For example, the traditional dichotomy into a written and spoken mode of communication (category 3) is an irrelevant division for a corpus of historical documents. Similarly, not every characteristic of the situation can be easily identified for every text in a historical corpus. The social background of the sender of a papyrus document and their personal relationships cannot always be described in detail. Part of the process of identifying relevant variables for the corpus in its current state, therefore, is to select a practical and suitable approach to the data. In this paper, I will describe the context along the following combinations of situational variables:

- (i) setting: date and context of use
- (ii) participants: nature and direction of interaction
- (iii) genre: communicative purpose and category
- (iv) production circumstances: method of production and stage of composition

All texts are divided into three main periods (Ptolemaic 300–1 BCE, Roman 1–300 CE and Byzantine 300–800 CE) in order to assess more general diachronic

¹³ Compare for example the categories of a context of situation by Firth (1950: 42–43), the eight components of speech defined by Hymes (1967) or the sociosemiotic interpretation of language along the three dimensions “field”, “mode” and “tenor” by Halliday, McIntosh & Stevens (1964).

differences. The setting also reflects the area of society in which the text is used, such as the private sphere or in the government administration.

The participants are classified according to their general role in society, the intended audience (or the lack thereof) and the direction of the interaction. Although the social and educational background of the participants involved is an important factor in explaining linguistic variation in papyri,¹⁴ there is no reliable way to establish the educational background and social status of all participants in the corpus based on language-external characteristics. Furthermore, the author or sender of the document needs to be distinguished from the scribe in historical documents. Since orthographic features are more likely to be a reflection of the social background and competence of the scribe than that of the author,¹⁵ the background of the scribe might be even more important than the social position of the author and addressee for the presence of orthographic variation. On the other hand, the intended register and general attitude towards standard orthography may also have been implied by the choices made by the author.¹⁶ These types of differences can hardly be quantified for the whole corpus, but some relevant distinctions can be made by studying documents in archives (§4).

Genre is vaguely defined by Biber (1988: 68) as “text categorizations made on the basis of external criteria relating to author/speaker purpose” and further specified as “text categories readily distinguished by mature speakers of a language”. This second notion is also reflected in the definition by Lee (2001: 46): “Genre is used when we view the text as a member of a category: a culturally recognised artifact, a grouping of texts according to some conventionally recognised criteria, a grouping according to purposive goals, culturally defined.” Genre in this study is based on the groups of texts generally distinguished in ancient society according to papyrological studies, combined into several larger categories with a similar communicative purpose (see §2.3.3).

Production circumstances are not always considered in register studies, but they seem to me of particular relevance for the papyrological corpus. Biber and Conrad (2009: 40) distinguish “real time / planned / scripted / revised and edited” and add that written registers are usually different from spoken registers in these respects (2009: 43–44). Just as in modern times, the production circumstances of a written document may have had an impact on the linguistic features. Whether the

¹⁴ Cf. Evans (2012).

¹⁵ Halla-aho (2018: 231–233).

¹⁶ Stolk (2019).

language of a particular papyrus has been produced freely at the spot, noted down from dictation, revised during a drafting process or copied from a written draft or model can make a significant difference to the type and amount of non-standard forms attested in a document.¹⁷ The importance of the method of production and stage of composition for orthographic variation will be examined in more detail in Section 4.

2.3 Categorization of documents

Texts within a corpus can only be distinguished based on extensive knowledge of the corpus itself. Overviews of the document types commonly distinguished by papyrologists can be found in Montevecchi (1973), Turner (1968: 127–153) and Palme (2009: 358–394). Montevecchi (1973: 86–89) gives a detailed overview of the contents of the corpus, but her organization into groups is partly thematic rather than functional. For example, the categories “administration”, “fiscal organization”, “jurisdiction” and “business, transportation, trade” identify groups of documents concerned with specific topics rather than reflecting different settings and communicative purposes. The categorization by Palme is more useful for linguistic approaches, since the categories are divided within several larger domains, such as “private documents”, “public life” and “interaction between state and individual”. These larger categories are partly reflected in the characteristics “setting” and “participants” (see below §2.3.1 and §2.3.2).

The basis for this categorization of all digitalized papyrus documents is the information about the contents of the text available at papyri.info and within Trismegistos.¹⁸ These content descriptions have been provided by the first editor (e.g. in the title of the edition), during entry of metadata in HGV and APIS, and/or by previous scholars working on specific genres within Trismegistos. Due to the extensive scale of the corpus it was not possible to consider every single text individually and the Greek text itself has only been consulted when the available metadata were insufficient or contradictory. Hence, this categorization relies to a large extent on the information that has been provided by

¹⁷ See Stolk (forthcoming), Stolk, Mihálykó & Grassien (forthcoming).

¹⁸ Paraliterary texts concerning religious and magical topics and school texts have been categorized, but they are not taken into account in this study. The same goes for very short texts such as name inscriptions on jars and mummy labels. Papyri without transcription and/or useful metadata and texts with uncertain contents are left out as well. In total, 45889 of the 51769 papyri with transcription in the database have been categorized into one of the seven genres and for 41977 of those the setting and participants could be determined as well.

previous scholars and aims only at a larger degree of unification and generalization at the level of the whole corpus.¹⁹

2.3.1 Setting

Traditionally, papyrus documents have been divided into private and public (or official) with respect to the context of use. This does not mean that these two domains were strictly separated in Greco-Roman Egypt, as government officials could produce both official and private documents and keep them together, such as the archive of Apollonios strategos (TM Archive ID 19). Private literate people would be in demand to take up public functions, especially from the later Roman period onwards.²⁰ Even though the same scribes and authors could use documents in both settings, the register requirements may have been different. Contracts can be produced for and used in more official or more private settings. Unless indicated otherwise, I assume for this paper that the majority of the preserved contracts originate from a private context of use rather than being solely produced for official archives. A private setting excludes texts that are used by the government, but it may contain documents from a professional, juridical or commercial context. The category of official documents includes documents related to public administration, justice and the military. Most of them were used at the level of the local administration (in the village or district capital), but few also relate to matters important to higher levels of the government.

2.3.2 Participants

Various types of relationships may exist between the people involved in the act of communication, such as between author and audience or sender and addressee. The author is taken to be the person in whose name the document is drawn up,

¹⁹ The basic level division into different genres (and subordinate topics) has been made available through the Trismegistos portal (see e.g. www.trismegistos.org/words). Previous mistakes may not have been noticed and the information gathered by different people could have been misleading in some cases. Documentary papyri are often fragmentarily preserved and their initial interpretation could be subject to revision. It is expected that the large amount of data and the generally abstract level of categorization will limit the consequences of possible misinterpretations. However, if anyone notices a mistake, please inform the author of this chapter or Trismegistos to help improving this resource.

²⁰ Ast (2015).

i.e. the sender of a letter or the acknowledging party of a contract, and the audience is the intended reader or addressee of the document. The relationships between the participants are defined with respect to the two different domains of society: private and official. These lead to four basic relationships and directions of communication:

- (1) interaction between officials
- (2) interaction between private persons
- (3) interaction from official to private
- (4) interaction from private to official
- (5) private administration
- (6) official administration

Documents in these last two groups are not primarily meant for communication but rather for the purposes of recording and archiving. They may or may not be read later by the scribe himself or his associates. Only the texts concerned with the interaction between private persons and private administration are expected to be produced in a private setting. Of course, there are many more relationships and more detailed social levels to be distinguished within private and official settings, such as positions in government, social classes and professions. This type of detailed information, however, is not readily available for the full corpus of documentary papyri.

2.3.3 Genre

The identification of genres is intended to stay close to the traditional divisions made by papyrologists to distinguish the different genres that existed in ancient society. As these labels have a tendency to become fuzzy, a prototype approach is applied to distinguish different levels of generalization (cf. Lee 2001: 48):

Table 2: Example of a prototype approach applied to genres of documentary papyri.

Superordinate		contract		declaration				
Basic level		contract	testament		petition		notification	
Subordinate		sale	lease	will	donatio mortis causa	enteuxis	petition to prefect	epicrisis declaration of death

An infinite amount of topics can be reflected in the lowest level of the individual instances, while the basic level distinguishes the genres as they are generally recognized in society.²¹ Document types which share a similar communicative function are combined into the seven superordinate genres. Inevitably, these genres bear a strong relationship with the setting in which they are commonly produced and the relations between the participants involved. The relations between the different variables have also been indicated in the characterisations below.

- 1) **Letters** are concerned with correspondence between private people or between government officials. Communication from private people to officials or vice versa often takes often a more specialized form, such as a declaration or pronouncement (see below). Orders for arrest and payment orders are also considered as a form of correspondence.
- 2) **Contracts** are produced as proof of a juridical agreement. Although they can be used in official settings, they usually concern a (juridical) relationship between private people. There are numerous different subtypes, such as a sale, loan, lease and marriage contracts, donations and testaments.
- 3) **Declarations** are letters to request or notify the government of (private) events. The largest group consists of petitions, but also other notifications to authorities, such as applications to *epicrisis* and *census*, notifications of birth, death, property returns, are included.
- 4) **Pronouncements** are formal announcements to inform government officials and citizens about rules and regulations. These pronouncements often originate in the higher level of the government and are only passed on through the lower levels of administration.
- 5) **Reports** are documents produced to collect and record information. Official registration was important in ancient society.²² These include council minutes and court proceedings as well as diaries of officials, land inspections, registers of official correspondence and abstracts of contracts to be kept in archives.
- 6) **Receipts** provide proof of delivery, payment or the execution of work. They form a continuum with contracts that can also be used to prove an exchange of money, such as an acknowledgement of debt or deposit, but they are generally shorter with fewer legal precautions. Documents to prove that work has been done, such as the *penthmeros* certificates, are also counted.

²¹ Taylor (1989: 48). Almost all papyrological texts apply some fixed formulas, but certain genres (such as contracts and certificates) are largely pre-composed, while others, such as letters, may contain more elements of free composition. As this study is focussed on the situational context, the direct linguistic context (specific terminology or fixed formulas) is not given special attention at this stage.

²² Palme (2009: 374–375).

- 7) **Lists** are mainly defined by their form as itemized collection of information. The documents in this group can be difficult to place in a social setting due to a lack of context.

3 Quantitative results

For this study, the presence of non-standard orthography as well as the total number of words per text is calculated in the Trismegistos Irregularities database. A maximum of one non-standard orthographic feature is counted per word. The normalized frequency (Nf), then, is the number of non-standard spellings per 1000 words for every text.²³ The normalized frequency is used as a dependent variable and the situational variables setting, participants and genre as independent variables in the following test. The results for each of these variables are presented and interpreted in Section 3.1, while other factors are considered in Section 3.2. The remaining situational variables, such as the social background of the scribe and the production circumstances will be addressed in Section 4.

3.1 Comparison of the situational variables

The average frequency of occurrence of non-standard orthographic features generally increases over time in the papyri from 14 per 1000 words in the Ptolemaic period to 31 per 1000 words in the Byzantine period. It is important to distinguish between the frequencies in the Ptolemaic and Byzantine period in this study, since the highest levels of non-standard features (16 per 1000 words) for any category during the Ptolemaic period are identical to the lowest average frequencies during the Byzantine period. These differences are highly significant ($p < 2.71 \times 10^{-157}$) in comparison with the other variables.²⁴ Table 3 presents the predictors for standard orthography in documentary papyri in Egypt in each of these three periods separately.²⁵

²³ In order to limit the effects of outliers from very short texts on the normalized frequency, only documents with a reliable word count and a minimum length of 10 words are taken into account.

²⁴ Based on linear regression analysis in a fixed effects model with the other independent variables.

²⁵ From the documentary texts for which all three situational variables could be determined (cf. fn. 18) with a minimum length of 10 words (cf. fn. 23), securely dated to the Ptolemaic (300–1 BCE), Roman (1–300 CE) or Byzantine (300–800 CE) period are counted (total 35024 texts).

Table 3: Linear regression analysis on three situational variables.²⁶

	Ptolemaic (300–1 BCE)			Roman (1–300 CE)			Byzantine (300–800 CE)		
Total texts (T)	6597			19004			9423		
Overall mean (Nf)	14			26			31		
	coef	T	Nf	coef	T	Nf	coef	T	Nf
Setting	p.<3.01e–05			p.<6.49e–103			p.<5.43e–25		
private	1.401	2772	15	7.507	5910	26	5.083	4461	36
official	–1.401	3825	13	–7.507	13094	21	–5.083	4962	26
Participants	p.<1.21e–16			p.<3.95e–192			p.<1.92e–14		
private people	4.157	3099	16	13.171	8067	36	5.690	6584	33
officials	–0.721	693	11	10.174	1044	33	4.975	614	32
private to official	3.693	1002	16	–1.727	2068	21	5.035	608	32
official to private	–2.205	987	10	–7.164	6067	16	0.161	877	27
private admin.	0.392	447	12	–6.570	394	16	–5.151	191	21
official admin.	–5.315	369	7	–7.885	1364	15	–10.710	549	16
Genre	p.<1.72e–07			p~0			p.< 5.74e–32		
letter	2.206	2282	15	23.998	3613	49	10.099	2749	38
contract	3.753	774	16	7.402	2552	33	–3.177	2314	25
pronouncement	–1.811	44	11	0.524	259	26	7.268	111	35
declaration	2.973	1002	16	–4.125	2069	21	3.954	608	32
receipt	–0.651	1682	12	–7.833	8756	18	3.695	2901	32
list	–2.733	667	10	–10.338	949	15	–9.623	531	18
report	–3.737	146	9	–9.628	806	16	–12.217	209	16

²⁶ The results in Table 3 are computed in R with the package Rbrul (see Johnson 2009). The absolute number of texts is given under T and the Nf shows the mean of the normalized frequency of the number of nonstandard words per 1000 words in those texts. The p-value per independent variable (p.) and the effect-size coefficients (coef) for each subcategory are added to show whether the effect of the variables is significant. A p-value < 0.05 is taken as a significant effect. The coefficients show the relative strength of the prediction between the subcategories. A larger coefficient (positive or negative) indicates a stronger effect than a smaller coefficient.

All situational variables have significant effects (see the p-values and effect size coefficients in Table 3) on predicting the number of non-standard features. A fixed effects model combining all three variables (log.likelihood -180948.7) shows that the variables genre ($p < 8.51e-84$) and participants ($p < 1.47e-05$) are the best predictors for orthographic variation in this corpus, while the setting does not provide a significant contribution to the other two.²⁷ Documents produced in a private setting generally contain a higher number of non-standard features than those from an official setting across all periods, but this difference is minimal in the Ptolemaic period (15 vs. 13 per 1000 words per text) and gets only more pronounced in the Byzantine period (36 vs. 26 per 1000 words per text).

As perhaps expected, texts produced as part of the interaction between private people contain a high number of non-standard features across all three periods (16, 36 and 33, respectively), while official administration produces the lowest numbers (with 7, 15 and 16 non-standard spellings per 1000 words). It should be noticed, though, that the frequency of non-standard spellings in documents directed from private to official is equally high during the Ptolemaic period (also 16 per 1000 words). During the Roman and Byzantine periods, the interaction between officials yields a particularly high number of non-standard features (33 and 32, respectively), not much below the average number of non-standard features in documents between private people.

These observations are confirmed by the results according to genre. Letters and contracts – which contain most of the interaction between private people and between officials – have the highest numbers of non-standard features in the Roman period (up to 49 per 1000 words). Administrative lists and reports tend to have the lowest number of non-standard features of all genres during these three periods, which is comparable to the low average frequencies found for documents used in private and official administration.

Combination of the situational variables allows us to specify in more detail in which situational context non-standard features appear most commonly. During the Ptolemaic period, contracts (Nf=16) and declarations (Nf=16) share the same high level of non-standard orthography as the private letters (Nf=16). During the Roman period, not only the letters between private people (Nf=56), but also contracts (Nf=33) and letters between officials (Nf=32) have a higher number of non-standard features than average. In the Byzantine period, letters

²⁷ The added value of testing these three predictors together in a multi-effect model is limited, since they are partly dependent on each other (see §2.3). A variable rule analysis of the three fixed effects together does not give any significant model for the Ptolemaic period.

(Nf=40) and receipts (Nf=34) between private people as well as pronouncements and nominations by officials (Nf=35) show the highest frequencies.

3.2 Writers and register

Although the combination of the variables participants and genre proves to be a reasonable predictor for the amount of orthographic variation in the corpus, this does not mean that this is the only or even most important explanation for the distribution of orthographic variation in papyri. Register-dependent linguistic features are typically identified in the lexical and grammatical domains of language rather than at the level of orthography.²⁸ It has been suggested that the use of non-standard orthography is mainly the result of the level of education of the individual scribe.²⁹ For example, Vierros (2012) has shown that there can be large differences in linguistic competence between scribes producing contracts in very similar situational contexts. Thus, the level of orthography achieved in these circumstances may be in the first place writer-dependent rather than register-dependent.

The importance of this factor for the general interpretation of the data can also be observed in the above quantitative results by comparing the standard deviation of each of the different genres. The high level of non-standard orthography in (private) letters, in particular, obscures a high level of variation within the category itself. While the standard deviation from the reported means lies between 26 and 36 for all other genres, letters show a standard deviation of 58 and letters between private people even 62. That means that private letters cannot straightforwardly be identified as a “low” register, even though it seems that an informal situational context coincides here with a high average number of non-standard orthographic features. Instead, different registers may need to be identified within this group of letters between private people. A drastic refinement of the variables covering the social background of the participants, including the scribe, seems necessary to explain whether the distribution of orthographic variation is dependent on the level of education of the scribe, the social status of the addressee or both (see also §4.1 and §4.2). Only a more detailed comparison of different types of texts written by the same scribe may reveal whether the variation is primarily related to *use* or to the *user* (see §4.1).

²⁸ Biber & Conrad (2009: 6).

²⁹ Evans (2012).

4 Registers in archives

The quantitative results show a relation between the occurrence of non-standard orthography and the situational context of the text (§3.1). Although general tendencies can be observed, there is a large degree of variation within some of the categories (§3.2). Archives can often provide more information about the role of individuals and the process of text production (cf. §2.2). Comparison of the results within and between several known archives allows us to address the role of the scribe and the production circumstances of the text in more detail and establish whether any of these other variables are relevant to explain the occurrence of orthographic variation in the corpus of documentary papyri. Below I examine the occurrence and distribution of non-standard orthography in several genres in the third century BCE Zenon archive (TM Archive ID 256), the second century BCE archive of the Katochoi of the Sarapieion (TM Archive ID 119), the first century CE archive of Kronion son of Apion, head of the *grapheion* of Tebtynis (TM Archive ID 93) and the sixth century CE archive of Dioscorus from Aphroditon (TM Archive ID 72).³⁰

4.1 Petitions, requests and lists in the Zenon archive

According to the results of the quantitative survey in Section 3, non-standard orthography seems most common in declarations, contracts and private letters during the Ptolemaic period. In the Zenon archive, dating to the mid third century BCE, all seven genres are represented and declarations have indeed the highest average frequency of non-standard features of all (Nf=23). The low level of orthography in the declarations in the Zenon archive could perhaps be explained by the nature of these requests and the roles of their participants. The majority (98 out of a 148 declarations) are in fact letters expressing a request or complaint to Zenon (or occasionally to one of the other protagonists of the archive). A well-known example is the petition of an Egyptian lady Senchons to Zenon (P.Mich. I 29) in which she complains that her she-ass has been taken and begs him to help her to get it returned. This Greek text has been painted with a brush by a native Egyptian scribe³¹ and contains one of the highest numbers of non-standard features in the group (Nf=189). These requests for help

³⁰ More information about these (and other archives) can be found through the Trismegistos portal, at www.trismegistos.org/arch/index.php, see also Vandorpe, Clarysse & Verreth (2015).

³¹ See Clarysse (1993: 196–199).

addressed to Zenon as the manager of the estate or just as a person in a higher position may not have required the same level of orthography as an official petition or *enteuxis* addressed to the king of Egypt. The wide variety of backgrounds of the writers of these petitions does not allow for a high average level of standardization and this was apparently not a major factor for the recognition of the submitted request. The linguistic variety in the requests to Zenon can be contrasted to the relative uniformity of the *enteuxeis* in the official archive of Diophanes, *strategos* of the Arsinoite district from the later third century BCE (TM Archive ID 80). Non-standard orthography occurs in this archive as well, but the frequencies are well below the average for the Ptolemaic period (Nf=8). Petitions to the king or a high official are also found in the Zenon archive, but the mere fact that they were found in the archive suggests that these were probably not official or final versions of the text.³² Some of them concern petitions by a third party that were copied (sometimes as part of a letter) and forwarded to Zenon or Apollonios. The general level of orthography in these copied petitions (Nf=26) is comparable to the requests addressed to Zenon (Nf=24).

Another small group contains drafts of petitions written in the name of Zenon. One might think that orthographic variation would be acceptable in such a preliminary version of a document, perhaps even expected to some extent, but the frequency of non-standard orthographic features in this group is the lowest of the petitions in the archive (Nf=12). That Zenon was a highly educated and conservative writer can also be observed from his other texts. When archaisms, such as *sandhi* and *crasis*, are left out, only very few non-standard spellings are found in the documents written in his own hand.³³ If anything can be concluded from this small group of autographs, it seems that non-standard orthographic features occur particularly in Zenon's accounts, personal notes and to-do-lists (e.g. P.Col. III 58; P.Cair. Zen. IV 59787). These may represent the situations in which standard orthography was least important to Zenon. This potentially register-based difference, however, is not reflected in the quantitative results of the papyrological corpus, where private administration, accounts and lists are categories attracting generally low levels of orthographic variation (see 3.1).

The Zenon papyri thus show the significance of the level of education of the writer for the occurrence of non-standard orthography. When a writer is able to produce standard orthography without difficulties, he is less likely to

³² Pestman (1981: 190–191).

³³ Clarysse (2009: 38–44).

produce non-standard features in any situational context. On the other hand, studying documents by the same writer – even a well-educated one – shows that the notion of register could be relevant to explain orthographic variation in different situations, albeit with slightly different results than attested in the full corpus written by numerous writers with various levels of competence.

4.2 Petitions and letters in the archive of the Katochoi of the Sarapieion

The highest average frequency of non-standard features in the Ptolemaic period (Nf=50), is found in the archive of the Katochoi of the Sarapieion³⁴ from the second century BCE. Bentein (2015a: 481) concludes that we can situate the documents in this archive on a register continuum from dreams and letters to petitions, where dreams contain “the largest amount of orthographical mistakes” and petitions the smallest. His ranking was based on a combination of linguistic features at various levels of the language, but the average frequencies of non-standard orthography (Nf=43 for declarations and Nf=58 for letters) give the same impression. The petitions addressed to the king (*enteuxeis*, Nf=20) seem to show a more standardized orthography than those addressed to a *strategos* or other official (Nf=53), but the genre and the status of the addressee are not the only factors responsible for the distribution of non-standard orthography in this archive.

The writer of the document and the version of the document at hand may also play a role, as briefly observed by Bentein (2015a: 469). In fact, when all petitions and letters in the archive are categorized based on the hand writing (following the identification by Wilcken in UPZ I), it becomes clear that the average frequency of non-standard features in the petitions (Nf=43) is mainly lowered by the petitions written by professional scribes in chancery hands (Nf=12), who are also responsible for most of the *enteuxeis*. The average number of non-standard features is much higher in the petitions written by Ptolemaios’ younger brother Apollonios (Nf=91). When we compare the orthography of the petitions and letters written by Apollonios, we find that Apollonios even produces more non-standard orthographical features in his petitions than in his letters (Nf=69). However, this should not automatically lead to the conclusion that Apollonios’ letters represent a “higher” register than his petitions. While the preserved letters by Apollonios are copies of official letters (UPZ I 23, I 26, I 37, I 38) and final versions of private

34 On this archive, see also Vierros (this volume), Bentein (this volume).

letters to his brother Ptolemaios (UPZ I 65, I 68, I 70, I 93), the majority of the petitions produced by Apollonios are preliminary drafts of petitions surviving in multiple copies (e.g. UPZ I 18, I 33, I 39). His copies of official letters by other scribes are more standardized (Nf=46) than the final versions of his own private letters (Nf=78), but his private letters seem again more standardized than these drafts of petitions (Nf=91). Hence, the level of education and experience of the writer (Apollonios or chancellery scribe), method of production (copy or draft) and the stage of composition of the text in question (preliminary or final product) count as important factors to explain the presence of non-standard orthographic features in this archive.

4.3 Contracts, lists and reports in the *grapheion* archive

The contracts from the archive of Apion and his son Kronion, heads of the record office (*grapheion*) of Tebtynis between 7 and 56 CE, have been used to show that “the majority of the professional scribes associated with the Tebtynis *grapheion* do not seem very concerned with the maintenance of standard orthography” (Bucking 2007: 237). This may be the case at a personal level, or perhaps even more generally among scribes at the lower levels of administration, as Bucking (2007) assumes, but what are the consequences of this for the levels of orthography generally achieved in juridical documents? Less focus on using standard orthography during scribal training and in scribal practice could explain the observed increase in non-standard forms in contracts during the Roman period. However, the average frequency of non-standard orthography in the contracts preserved in the *grapheion* archive (Nf=99) is well above the average for contracts during that period and this may be related to the stage of production of these contracts.

The village record office had two main tasks: to compose contracts for their private customers and to keep a register of all produced contracts for the official administration. For the first stage of the production of a new contract, the contracting parties convened at the *grapheion* and a contract was drawn up by one of the scribes from the office.³⁵ After that, the contracting parties added their subscription to the bottom half of several papyrus sheets in order to provide each party with its own copy of the whole contract at a later stage. The parties subscribed the contract themselves, if needed with help from a relative or one

35 For the procedure see Husselman et al. (1944: 3–11).

of the scribes often called upon.³⁶ The majority of the contracts preserved in the archive comprise in fact these subscriptions on partial contracts that still needed to be completed and collected by the customer. Not only the fact that these particular documents were never finished, but especially the part of the document that is preserved may explain the frequent occurrences of non-standard orthography. When compared, the average frequency of non-standard spellings (Nf=114) is much higher in the documents that contain only the subscriptions of the parties than in the other contracts in which also (part of) the body of the document is preserved (Nf=69). Subscriptions to contracts are generally produced by people with various levels of (orthographic) competence.³⁷ The fact that a signature was supposed to be written in the hand of the party itself rather than added by someone else was clearly considered more important than the correct use of orthography in this part of the document. The abnormally high ratio of partial documents and papyri containing subscriptions only contributes to the high frequency of non-standard features in the *grapheion* archive.

Apart from producing the contracts for their customers, the *grapheion* also had to forward all contracts to the central administration (in a τόμος συγκολλήσιμος ‘pasted roll’), compile a document with abstracts to these contracts (εἰρόμενον) and provide a list of all contracts by title (ἀναγραφή). Parts of the latter two types of registers have been found in the archive as well. The average frequency of non-standard forms in these reports (Nf=72) is comparable to the numbers encountered in the more complete versions of the contracts.³⁸ However, the preserved lists and reports represent again different stages in the production of the final report. As Husselman (1970: 237) illustrates, preliminary entries of various types were made by Kronion in P.Mich. II 128, to be transferred later to P.Mich. II 123 and P.Mich. V 238, and again copied and put in a chronological order in P.Mich. V 240. During the compilation of the final report, the contents were checked and adjusted where necessary, but also minor changes were made to the orthography. For example, in the draft entries by Kronion in P.Mich. II 128, ii, ll. 19, 23 and 25, the words γεργίας (instead of γεωργίας) ‘of cultivation’, κομογρ(αμματέως) (instead of κωμογραμματέως) ‘of a village scribe’ and ἐνυκήσεω(ς) (instead of ἐνοικήσεω(ς)) ‘of lodging’, seem reproduced in identical (non-standard) spellings in the chronological list in P.Mich. V 240, ll. 14, 17 and 19, while the words

³⁶ See Husselman et al. (1944: 21–22).

³⁷ See Youtie (1971).

³⁸ Based on the published reports listed by Husselman (1970: 227–230).

ἀλοπολῶ(ν) ‘of salt merchants’, περαχορή(σεως) ‘of cession’ and φυνικό(νος) (instead of φοινικῶνος) ‘of a palm-grove’ in P.Mich. II 128, iii, ll. 10, 13, 22 become (partially) corrected to ἀλοπω(λῶν), παραχο(ρήσεως) and φυνικῶνος(ς) in P.Mich. V 240, ll. 31, 33 and 42.

The frequency of non-standard spellings in the registers and accounts of expenses of the *grapheion* is mainly the result of a limited set of lexemes for which the supposedly non-standard orthographic variant is used consistently.³⁹ P.Mich. II 122 contains a model for the registration of contracts which Kronion may have produced as an example for another scribe in the office.⁴⁰ In this document all names are replaced by the indefinite pronouns, such as ὁ δεῖνα τοῦ δεινοῦ ‘NN son of NN’, and these are almost consistently spelled as ὁ δῖνα τοῦ δῖνος or τοῦ δῖνατος. This word on its own makes up the majority of the non-standard spellings in this text.

Furthermore, the high number of abbreviations in registers and accounts lowers the potential number of candidates for non-standard spellings. As more than half of the words in these registers are abbreviated, the number of non-standard spellings pertaining to morphemes (often word-final) is reduced to 5 per 1000 words, compared to 23 in the contracts from the archive. The same phenomenon applies to all lists and reports in the Roman period. While letters, contracts and receipts have non-standard spellings exclusively in morphemes in respectively 10, 6 and 5 words out of a 1000, the lists and reports count only 3 and 2, respectively. The repetitive nature of the registers and accounts combined with their relatively limited vocabulary and fixed syntactic structures should make the orthography of these genres easier for scribes to master and potentially reduce the number of non-standard orthographic features compared to contracts and letters. On the other hand, as observed in the *grapheion* archive above, when a certain non-standard spelling of a frequent word is used repeatedly in the same text, they could still add up to large numbers for single documents.

4.4 Pronouncements and petitions in the Dioscorus archive

Pronouncements in the form of edicts by the Roman emperor or the prefect of Egypt are rarely transmitted to us in their original version. When they surface

³⁹ See also Bucking (2007: 237, Table 3) for the variation between πρόκειται and πρόκειται in this archive.

⁴⁰ Cf. Bucking (2007: 237, Table 4).

on a papyrus in a village in Egypt, the orthography cannot straightforwardly be taken to reflect the standards of the imperial chancery. This has been shown by Bucking (2007: 233) for the letter of the emperor Claudius to the Alexandrians in P.Lond. VI 1912 (see also §1.2) and the same can be assumed for many other examples. P.Yale I 61 contains an edict by Soubatianus Aquila, the prefect of Egypt between 206 and 210 CE, announced by Sarapion, the *strategos* of the Arsinoite district. The text is written in an impressive chancery hand, but contains several non-standard orthographic forms. The contents, morphology and syntax have been revised by the same hand as the one who approved of the whole text, probably the *strategos* himself. This procedure of (re)composing, copying and distributing important messages throughout the provinces, inevitably led to linguistic variation between the different versions of the text.

During the Byzantine period, the procedure to compose imperial documents can be observed in more detail in the archive of Dioscorus of Aphrodito from the middle of the sixth century CE. The imperial rescripts P.Cair.Masp. I 67024–67025, 67026–67027 and 67028 are drafts written by Dioscorus himself and his companion.⁴¹ During their visit to Constantinople, they produced their own rescripts, i.e. answers to their own requests.⁴² Several versions of these rescripts have been preserved in the archive, such as the reformulation of 67024 recto on the verso by Dioscorus, copied and revised again by his companion in 67025. The various stages of drafting of the rescripts show a high level of skill. A generally high level of orthographic standardization is visible in these drafts and revision is mainly aimed at stylistic refinement.

While these drafts for imperial rescripts left very little to be desired from an orthographic point of view, the same cannot be said of all petitions in Dioscorus' archive. Some petitions contain particularly high frequencies of non-standard spellings, for example P.Cair.Masp. I 67002, 67006, 67020 and P.Lond. V 1674. Their apparently lower orthographic standards co-occur with a high frequency of scribal corrections. P.Cair.Masp. I 67002 and P.Lond. V 1674 were written by Dioscorus himself⁴³ and are likely to be drafts.⁴⁴ The corrections consist almost entirely of superlinear additions of words and short phrases. Hardly any orthographic or morphological corrections are made, while there are several possible candidates, e.g. ἐπί 'upon' for ἐπεῖ 'when' in margin of P.Lond. V 1674, l. 21, ἐπράϊτευσεν for ἐπράϊδευσεν 'he plundered' in P.Cair.Masp. I 67002, ii, l. 24 (cf. πραιδέουσι in iii 13) and ὑπολέλιπται for ὑπολέλειπται 'is left' in 1674, l. 94

⁴¹ Perhaps his cousin Dioscorus, see Van Minnen (2003), Zuckerman (2004).

⁴² Van Minnen (2003), Feissel (2004).

⁴³ See Del Corso (2008).

⁴⁴ Fournet & Gascou (2004: 145–146, 168–169), Keenan (2008: 173).

(cf. ὑπο|λέλειπται in 67002, iii, ll. 11–12). P.Cair.Masp. I 67006 and 67020 were not written by Dioscorus himself, but he did make corrections to these texts.⁴⁵ Orthographic corrections are more frequent than stylistic corrections in these two petitions, especially in 67020, which Fournet and Gascoü (2004: 153) assume to have been dictated to the scribe by Dioscorus. Not only omitted vowels and consonants are inserted by Dioscorus in this document, but also interchanged consonants and itacisms are corrected this time, e.g. ἐπεῖ ‘when’ corrected to ἐπί ‘upon’ in l. 3, τοτῆρες corrected to δοτῆρες ‘givers’ in l. 12, χαλκῖς corrected to χαλκεῖς ‘smiths’ in l. 17. While non-standard orthography does not seem to be the focus of attention in preliminary drafts, Dioscorus did make orthographic corrections to copies of more complete petitions produced by his scribes.⁴⁶ Thus, the scribe, method of production and stage of composition of these documents is highly relevant to explain the presence or absence of orthographic variation.

5 Conclusions and discussion

In this paper, I set out to examine the relationship between the use of non-standard orthography and the situational context in the corpus of Greek documentary papyri from Egypt. Based on quantitative testing, non-standard orthography generally seems to occur in documents related to the interaction between private people, such as letters and contracts. It is also frequently encountered in declarations to the government during the Ptolemaic and Byzantine periods, letters between officials in Roman period and pronouncements and nominations by officials in the Byzantine period. The lowest frequencies of non-standard orthography are found in lists and reports in private and official administration. The distinction between private and official settings is therefore not an essential parameter to explain the distribution of non-standard orthography, but the participants and genre are relevant factors. Some (linguistic) characteristics of the genres themselves may help to explain these differences. The fixed structure, repetitive vocabulary and frequent use of abbreviations could potentially reduce the number of non-standard orthographic features in lists and reports. On the other hand, the registers and accounts of expenses in the *grapheion* archive show that a repetitive use of certain non-standard forms, especially in a (short) text with a limited range of vocabulary, could also easily increase the normalized frequency of non-standard features. Even though contracts

⁴⁵ Fournet & Gascoü (2004: 148, 153).

⁴⁶ See also Stolk (forthcoming).

also contain many fixed formulas, non-standard orthography is more common in this genre than in others during the Ptolemaic and Roman periods. This may be related to the way in which contracts and their subscriptions are produced, the level of education of the writers involved and, perhaps, the general attitude towards standard orthography in scribal offices.

Although the setting, the participants and the genre may explain part of the distribution of orthographic variation, closer analysis reveals several other important factors. The particularly high level non-standard orthography in private letters across all periods co-occurs with a high level of variation within the group itself. Because levels of education could differ greatly between individuals, the level of orthography is always closely related to the choice of scribe. The variety of backgrounds of the writers of requests to Zenon shows why these documents generally have a lower standard than for example official *enteuxeis* to the king during the same period. A more experienced scribe could be employed to increase the level of orthography in petitions, as was shown by the differences between the petitions written by Apollonios and those written by professional scribes in Ptolemaios' archive. Official contracts are typically produced by multiple writers and the different levels of experience may also create a difference between the orthographic variation in the body of the text and in the subscriptions, as shown in the partially completed documents in the *grapheion* archive. In documents written by the same writer, register-based variation can be identified more clearly. The highly-educated secretary Zenon produces standard spelling in almost every situation, but this attention to detail may drop slightly in lists and notes for personal use. The well-educated notary Dioscorus also manages to produce a higher standard of orthography in imperial rescripts than in his other genres.

Since many scribes had to make an effort to produce a document in standard orthography, the method of production and version of the text that is preserved to us seem important variables for orthographic variation. Apollonios produces more standard orthography in his copies of letters by other scribes than in his own letters, while his final versions of private letters are again more standardized than his drafts of petitions. The amount of linguistic variation, especially at the level of orthography, could increase or decrease during the production process. Kronion shows that orthography can be improved in the process from draft to the final version and Dioscorus makes multiple orthographic corrections to documents, potentially increasing the level of orthographic standardization. When documents produced at the highest levels of society, such as official texts sent to and from Alexandria or imperial correspondence, surface in little villages in the Egyptian countryside, they are most likely versions handed down to us through several stages of copying by various scribes, possibly decreasing the level of standardization. Still, the

high standards pursued by Dioscorus by means of corrections in his drafts and in the documents produced by his scribes show that standard orthography was still highly valued in rescripts and petitions, despite their sometimes different appearance in the versions preserved to us.

Apart from the register expectations related to the genre of the document and the participants involved, the choice of scribe and the process of textual production thus seem important factors to explain the presence of non-standard orthographic features in the corpus of documentary papyri. It would be advisable, therefore, to take account of the method of production and stage of composition of the text in question in future studies of linguistic variation.⁴⁷

References

- Papyrus editions are cited according to J. Oates et al. Checklist of Editions of Greek, Latin, Demotic, and Coptic Papyri, Ostraca, and Tablets www.papyri.info/docs/checklist (accessed 22 December 2017).
- Ast, Rodney. 2015. Writing and the city in later Roman Egypt: Towards a social history of the ancient scribe. *CHS Research Bulletin* 4.1 <nrs.harvard.edu/urn-3:hInc.essay:AstR.Writing_in_the_City_in_Later_Roman_Egypt.2016>.
- Bentein, Klaas. 2012. Perfect periphrases in Post-classical and Early Byzantine Greek: An ecological-evolutionary account. *Journal of Greek Linguistics* 12. 205–275.
- Bentein, Klaas. 2013. Register and the diachrony of Post-classical and Early Byzantine Greek. *Revue Belge de Philologie et d'Histoire* 91. 5–44.
- Bentein, Klaas. 2015a. The Greek documentary papyri as a linguistically heterogeneous corpus: The case of the katochoi of the Sarapieion-archive. *Classical World* 108. 461–484.
- Bentein, Klaas. 2015b. Particle-usage in documentary papyri (I–IV AD): An integrated, sociolinguistically-informed approach. *Greek Roman and Byzantine Studies* 55. 721–753.
- Bentein, Klaas. 2015c. Minor complementation patterns in Post-classical Greek (I–VI AD): A socio-historical analysis of a corpus of documentary papyri. *Symbolae Osloenses* 89. 104–147.
- Bentein, Klaas. 2017. Finite vs. non-finite complementation in Post-classical and Early Byzantine Greek: Towards a pragmatic restructuring of the complementation system? *Journal of Greek Linguistics* 17. 3–36.
- Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, Douglas. 1995. *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.

⁴⁷ My research was funded by the Research Foundation - Flanders (FWO), at Ghent University, and The Research Council of Norway (NFR COFUND), at the University of Oslo.

- Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.
- Bucking, Scott. 2007. On the training of documentary scribes in Roman, Byzantine, and early Islamic Egypt: A contextualized assessment of the Greek evidence. *Zeitschrift für Papyrologie und Epigraphik* 159. 229–247.
- Clarysse, Willy. 1993. Egyptian scribes writing Greek. *Chronique d'Égypte* 68. 186–201.
- Clarysse, Willy. 2009. The Zenon papyri thirty years on. In Guido Bastianini & Angelo Casanova (eds.), *100 Anni di Istituzioni Fiorentine per la Papirologia: Atti del convegno internazionale di studi Firenze, 12–13 giugno 2008*, 31–43. Firenze: Istituto papirologico G. Vitelli.
- Clarysse, Willy. 2010. Linguistic diversity in the archive of the engineers Kleon and Theodoros. In Trevor V. Evans & Dirk D. Obbink (eds.), *The language of the papyri*, 35–50. Oxford: Oxford University Press.
- Depauw, Mark & Joanne Vera Stolk. 2015. Linguistic variation in Greek papyri: Towards a new tool for quantitative study. *Greek, Roman & Byzantine Studies* 55. 196–220.
- Del Corso, Lucio. 2008. Le scritture di Dioscoro. In Jean-Luc Fournet & Caroline Magdelaine (eds.), *Les archives de Dioscore d'Aphrodité cent ans après leur découverte*, 89–115. Paris: de Boccard.
- Evans, Trevor V. 2012. Linguistic and stylistic variation in the Zenon Archive. In Martti Leiwo, Hilla Halla-aho & Marja Vierros (eds.), *Variation and change in Greek and Latin*, 25–42. Helsinki: Suomen Ateenan-Instituutin säätiö.
- Feissel, Denis. 2004. Pétitions aux empereurs et formes du rescrit dans les sources documentaires du IV^e au VI^e siècle. In Denis Feissel & Jacques Gascou (eds.), *La pétition à byzance*, 33–52. Paris: Association des Amis du Centre de recherche d'histoire et civilisation de Byzance.
- Firth, John Rupert. 1950. Personality and language in society. *The Sociological Review* 42. 37–52.
- Fournet, Jean-Luc & Jacques Gascou. 2004. Liste des pétitions sur papyrus des V^e–VII^e siècles. In Denis Feissel & Jacques Gascou (eds.), *La pétition à byzance*, 141–196. Paris: Association des Amis du Centre de recherche d'histoire et civilisation de Byzance.
- Gignac, Francis T. 1976. *A grammar of the Greek papyri of the Roman and Byzantine periods*. Vol. 1: *Phonology*. Milano: La Goliardica.
- Halla-aho, Hilla. 2010. Linguistic varieties and language level in Latin non-literary letters. In Trevor V. Evans & Dirk D. Obbink (eds.), *The language of the papyri*, 171–183. Oxford: Oxford University Press.
- Halla-aho, Hilla. 2018. Scribes in private letter writing: Linguistic perspectives. In Jennifer Cromwell & Eitan Grossman (eds.), *Scribal repertoires in Egypt from the New Kingdom to the Early Islamic Period*, 227–239. Oxford: Oxford University Press.
- Halliday, Michael A. K., Angus McIntosh & Peter Stevens. 1964. *The linguistic sciences and language teaching*. London: Longman.
- Hickey, Raymond. 2010. Linguistic evaluation of earlier texts. In Raymond Hickey (ed.), *Varieties of English in writing: The written word as linguistic evidence*, 1–14. Amsterdam: Benjamins.
- Husselman, Elinor M., Arthur E.R. Boak & William F. Edgerton. 1944. *Papyri from Tebtunis II (Michigan papyri V)*. Ann Arbor: University of Michigan Press.

- Husselman, Eleanor M. 1970. Procedures of the record office of Tebtunis in the first century A.D. In Deborah Hobson Samuel (ed.), *Proceedings of the 12th International Congress of Papyrology, Ann Arbor, 13–17 August 1968*, 223–238. Toronto: Hakkert.
- Hymes, Dell. 1967. Models of the interaction of language and social setting. *Journal of Social Issues* 23. 8–28.
- James, Patrick. 2007. *Retention and retreat: Complementary participles and infinitives with verbs of perception and declaration in the Roman and Byzantine documentary papyri*. PhD dissertation, University of Cambridge.
- Johnson, Daniel E. 2009. Getting off the GoldVarb Standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3. 359–383.
- Keenan, James G. 2008. 'Tormented Voices': P.Cair.Masp. I 67002. In Jean-Luc Fournet & Caroline Magdelaine (eds.), *Les archives de Dioscore d'Aphrodité cent ans après leur découverte*, 171–180. Paris: de Boccard.
- Lee, David Y. W. 2001. Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology* 5(3). 37–72.
- Milroy, James & Lesley Milroy. 1985. *Authority in language: Investigating language prescription and standardisation*. London: Routledge & Kegan Paul.
- Montevocchi, Orsolina. 1973. *La papirologia*. Torino: Società editrice internazionale.
- Palme, Bernhard. 2009. The range of documentary texts: Types and categories. In Roger S. Bagnall (ed.), *The Oxford handbook of papyrology*, 358–394. Oxford: Oxford University Press.
- Pestman, Pieter W. 1981. *A guide to the Zenon archive (P. L.Bat. XXI)*. Leiden: Brill.
- Porter, Stanley E. & Matthew B. O'Donnell. 2010. Building and examining linguistic phenomena in a corpus of representative papyri. In Trevor V. Evans & Dirk D. Obbink (eds.), *The language of the papyri*, 287–311. Oxford: Oxford University Press.
- Rutkowska, Hanna & Paul Rössler. 2012. Orthographic variables. In Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre (eds.), *The handbook of historical sociolinguistics*, 221–244. Chichester: Wiley-Blackwell.
- Stolk, Joanne Vera. 2018. Encoding linguistic variation in Greek documentary papyri: The past, present and future of editorial regularization. In Nicola Reggiani (ed.), *Digital papyrology II: Case studies on the digital edition of Ancient Greek papyri*, 119–137. Berlin: de Gruyter.
- Stolk, Joanne Vera. 2019. Itacism from Zenon to Dioscoros: Scribal corrections of <ι> and <ει> in Greek papyri. In Sofia Torallas Tovar & Alberto Nodar (eds.), *Proceedings of the 28th International Congress of Papyrology, Barcelona, 1–6 August 2016*, 690–697. Barcelona: Publicacions de l'Abadia de Monsterrat, Universitat Pompeu Fabra.
- Stolk, Joanne Vera. Forthcoming. Scribal revision in the process of text production: A linguistic typology of scribal corrections in Greek documentary papyri. In Martti Leiwo, Sonja Dahlgren, Hilla Halla-aho and Marja Vierros (eds.), *Act of the scribe: Interfaces between scribal work and language use: A workshop, The Finnish Institute at Athens, April 6–8, 2017*.
- Stolk, Joanne Vera, Ágnes Mihálykó & Céline Grassien. Forthcoming. The language of the liturgy: Greek used by Theban scribes. In Anastasia Maravela & Ágnes Mihálykó (eds.), *New perspectives on religion, education, and culture at Christian western Thebes (VI–VIII)*.
- Taylor, John R. 1989. *Linguistic categorisation: Prototypes in linguistic theory*. Oxford: Clarendon Press.

- Teodorsson, Sven-Tage. 1977. *The phonology of Ptolemaic Koine*. Göteborg: Acta Universitatis Gothoburgensis.
- Turner, Eric G. 1968. *Greek papyri: An introduction*. Oxford: Clarendon Press.
- Vandorpe, Katelijn, Willy Clarysse & Herbert Verreth. 2015. *Graeco-Roman archives from the Fayum*. Leuven: Peeters.
- Van Minnen, Peter. 2003. Dioscorus and the law. In A.A. MacDonald, M.W. Twomey & G.J. Reinink (eds.), *Learned antiquity scholarship and society in the Near East, the Greco-Roman world, and the Early Medieval West*, 115–133. Leuven: Peeters.
- Vierros, Marja. 2012. *Bilingual notaries in Hellenistic Egypt: A study of Greek as a second language*. Brussel: Koninklijke Vlaamse Akademie van België voor Wetenschappen en Kunsten.
- Youtie, Herbert C. 1971. Βραδέως γράφων: Between Literacy and Illiteracy. *Greek, Roman & Byzantine Studies* 12. 239–61.
- Zuckerman, Constantin, 2004. *Les deux Dioscore d'Aphrodité ou les limites de la pétition*. In Denis Feissel & Jacques Gascou (eds.), *La pétition à byzance*, 75–92. Paris: Association des Amis du Centre de recherche d'histoire et civilisation de Byzance.