

Fabian Barteld, Stefan Hartmann and Renata Szczepaniak*

The usage and spread of sentence-internal capitalization in Early New High German: A multifactorial approach

DOI 10.1515/flin-2016-0015

Submitted July 6, 2015; Revision invited December 12, 2015;

Revision received April 1, 2016; Accepted May 31, 2016

Abstract: Sentence-internal capitalization of nouns is a characteristic of written Standard German. The sixteenth and seventeenth centuries have been identified as the crucial period for the development of this graphemic convention. Previous studies have shown that animacy played a major role in the spread of sentence-internal capitalization. On the basis of the transregional SiGS-corpus consisting of 18 protocols of witch trials (hand-)written between 1588 and 1630, we propose word frequency as an additional factor and test for its interaction with animacy. Our data reveal that the proportion of capitalized words denoting humans and animate concepts increases rapidly, while the capitalization of lexemes referring to concrete and abstract concepts remains stable at a lower level. A binomial mixed-effects model shows a highly significant effect of frequency and a significant interaction between frequency and animacy. In sum, our data show how cognitive, pragmatic, and usage factors conspire in the gradual emergence of a graphemic convention. We therefore argue that the previously neglected graphemic dimension can add important insights to an empirically based theory of the language-cognition interface.

Keywords: sentence-internal capitalization, Early New High German, animacy, frequency effects, handwritten texts

1 Introduction

In the standard orthography of modern German, sentence-internal capitalization is used not only to mark proper names, titles of honor, and (address) pronouns,

*Corresponding author: **Renata Szczepaniak**, Universität Hamburg, Von-Melle-Park 6, 20146 Hamburg, Germany, E-mail: renata.szczepaniak@uni-hamburg.de

Fabian Barteld: E-mail: fabian.barteld@uni-hamburg.de, **Stefan Hartmann:** E-mail: stefan.hartmann@uni-hamburg.de, Universität Hamburg, Von-Melle-Park 6, 20146 Hamburg, Germany

but to highlight every noun or nominalization that is the head of a noun phrase. Examples include *das groß-e Haus* ‘the big house’, and also *das groß-e Aber* ‘the big but’ (see Munske 1995a, 1995b; Eisenberg 2013). It has been argued that sentence-internal capital letters support and accelerate the decoding process by emphasizing the structure of the noun phrase (see Bock et al. 1989; Gfroerer et al. 1989, among others).

The crucial period for the development of this reader-friendly graphemic convention is Early New High German (ENHG; 1450–1750). This language period is characterized by an intensified use of writing as a medium for official and private communication. The increasing use of written language was supported by at least two phenomena: the invention of moveable-type printing in the middle of the fifteenth century and the emergence of the modern state. The former led to a steadily increasing number of printed books (see Buringh and Van Zanden 2009) as it made the reproduction and distribution of reading materials easier and cheaper. The development of the modern state entailed a certain degree of bureaucratization which led to an intensified use of (hand) writing and consequently to an increased need for literate citizens. While the literacy rate was low at the end of the fifteenth century (between 1% and 4% of the total population), the number of readers and writers grew substantially from the seventeenth century onwards as general compulsory education was introduced. In the middle of the eighteenth century, the literacy rate of the rural population reached 60% (Gauger 1994; Knoop 1994).

According to Maas (1995a: 90), the spread of sentence-internal capitalization can partly be attributed to the economic interests of book printers. Printers aimed to make their books more attractive, which also entailed making the texts more easily comprehensible for the reading public. According to the seventeenth-century grammarian Johann Rudolph Sattler, book printers used capital letters so that “the simple-minded can better understand it i. e., the text]” (“und könne es der einfeltige desta besser verstehen”, Sattler 1975[1617]: 17–18). Consequently, ever more printers made use of sentence-internal capitalization (see Bergmann and Nerius 1998; Rössler 1998; Bergmann 1999).

It is important to note that the use of sentence-internal capital letters did not start with printed texts since the tendency to highlight words by means of capitalization had already been observable in handwritten texts created before the invention of moveable-type printing. In Old High German (750–1050 AD), there is a tendency toward capitalization especially for proper nouns (see Weber 1958; Labs-Ehlert 1993). After Gutenberg’s invention, however, the spread of sentence-internal capitalization was faster in printed books than in handwritten texts (see Moulin 1990). The use of sentence-internal capitalization

in private writing was only established after the introduction of general compulsory education starting in the seventeenth century (Maas 1995a: 91) (see Section 2.3).

In this paper, we will analyze the occurrence of sentence-internal capital letters in the SiGS corpus,¹ which consists of handwritten protocols of witch trials from the sixteenth and seventeenth centuries with official status (Topalović 2003; Szczepaniak and Barteld 2016). Our main goal is to determine the factors responsible for the use of capital letters in this corpus and to compare them with previous studies discussing the distribution and historical spread of capitalization in printed books and handwritten texts.

Previous research has identified two factors in the spread of sentence-internal capitalization. One determining factor is of a pragmatic nature, as capitalization serves as a means for emphasis and deference. For instance, Bergmann and Nerius (1998) have shown that *Nomina sacra*, i. e., words referring to deities, saints, etc., and honorific titles are capitalized very early, which indicates that capitalization was used for marking deference toward religious entities and people with a higher social status. Second, as mentioned above, sentence-internal capitalization has been accounted for as a reader-friendly strategy that facilitates decoding. In this regard, strategies related to textual content, prosody, semantics, morphology, lexicon, and syntax have been identified (see Maas 1992; Fuhrhop 2009; Eisenberg 2013, among others). Previous research has focused on one semantic strategy in particular: as we will discuss in more detail in Section 2, an important factor in the spread of sentence-internal capitalization is animacy. As an animacy-based capitalization strategy highlights the cognitively salient concepts of humans and other living beings, it seems reasonable to assume that it can indeed support the decoding process (see Dahl and Fraurud 1996; Dahl 2008; Szczepaniak 2011; Dietrich and van Nice n.d.).

Another potential factor that has only played a minor role in diachronic studies so far but has figured prominently in synchronic psycholinguistic research is the graphemic stability and uniformity of a word: it supports a word's recognizability and thus accelerates the decoding process (see Section 3). As the graphemic system of German became increasingly stable from the seventeenth century onwards (see Voeste 2008), it can be expected

¹ This corpus was created within the framework of the joint project "Development of Sentence-internal Capitalization in German" (*Entwicklung der satzinternen Großschreibung im Deutschen*, SiGS for short), funded by the German Research Foundation (DFG) in 2013–2014 (SZ 280/2-1 and KO 909/12-1); see www.uni-muenster.de/SIGS-Projekt.

that this process of stabilization affected the spread of sentence-internal capitalization as well. More specifically, it is hypothesized that high-frequency words tend to stabilize their *gestalt* faster than low-frequency words and that the trend toward capitalization will therefore be more pronounced for the former than for the latter.

To test this hypothesis, we will examine the influence of animacy and frequency on the development of sentence-internal capitalization as well as their interaction in our corpus of handwritten witch trial protocols. To this purpose, we will consider the behavior of individual lexemes. Furthermore, we will take the influence of different scribes into account.

While previous research on the phenomenon in question was mainly based on printed texts, our study investigates the spread of capitalization in handwritten texts. Although they were written for official purposes, the witch trial protocols in our corpus were produced with considerably less planning effort than printed books. Thus, they exhibit a higher degree of spontaneity, especially so in the case of immediate transcripts. However, some texts in our corpus are copies of transcripts that might have been altered in the copying process; this makes it all the more important to take the influence of individual scribes into account in the analysis.

This leads us to a related advantage of handwritten texts as compared to printed books in the investigation of graphemic phenomena. While the preparation of printed texts often included a number of different people (see, e.g., Kaempfert 1980), the witch trial protocols were, in most cases, produced by one individual. In some rare cases, even the name of the scribe is given in the data. In Section 5.3, we will discuss the influence of idiolectal factors for graphemic variation in more detail.

In the remainder of this paper, we will first present a brief overview of previous research on the emergence and spread of sentence-internal capitalization (Section 2) and on the role of frequency in language change (Section 3). Then, we will introduce the corpus of protocols used for the present study (Section 4). In Section 5, we will discuss the individual factors that we hypothesize to play a role in the use of sentence-internal capitalization in our corpus data. We then use binomial logistic mixed-effects modeling to assess the influence of these individual factors as well as potential interactions between them (Section 6). In the final section, we summarize our results and discuss the possibilities and limitations of using mixed-effects modeling for investigating graphemic changes. In addition, we outline potential avenues for future research.

2 The development of sentence-internal capitalization in prints and handwritten texts: The influence of the animacy scale

While graphemics, and graphemic change in particular, can overall be considered a neglected topic in linguistics, there has already been a good amount of research on the spread of capitalization in the German writing system, drawing both on printed and on handwritten texts. In this section, we will give a brief sketch of this research. We will focus on the sentence-internal capitalization of nouns and introduce factors supporting capitalization as identified in the previous studies.

The development of sentence-internal capitalization in German printed texts has been shown to be influenced by various factors (see Bergmann and Nerius 1998). In their corpus comprising texts from 1500 to 1710, Bergmann and Nerius (1998) observe a spread of capitalization that, we suggest, resembles a grammaticalization process where capitalization starts with a pragmatically driven usage, spreads within nouns along the animacy scale (see Section 5.1), and finally leads to the development of a syntactic head-marking rule (see also Szczepaniak 2011).

The data investigated by Bergmann and Nerius (1998) show that the use of uppercase letters in the sixteenth century was determined by pragmatic factors in the broadest sense.² Official titles marking individuals at the top of the social, political, or administrative hierarchy were capitalized, as were *Nomina sacra* such as *god* or *Jesus*. At the same time, the capitalization of proper nouns, which had already started in handwritten texts (see Weber 1958; Labs-Ehlert 1993), suggests that, apart from the (socio-)pragmatic factor of reverence, the cognitive category of individuality played a key role in the use of capitalization. Importantly, capitalization of anthroponyms preceded that of toponyms (see Labs-Ehlert 1993). As it seems reasonable to assume that persons are conceptualized as more individuated than places, this lends further support to the idea that factors like individuality and animacy determine a printer's choice for or against capitalizing a specific word.

Note that individuality can be seen as corresponding with animacy (see, e. g., Yamamoto 2008). Individuation, i. e., the conceptualization of an entity as an individual, is the more likely the higher the entity's position is on the

² Note that the use of capitals determined by pragmatic factors was not limited to nouns. In the present paper, however, we will focus exclusively on nouns; for a discussion of the rise and fall of capitalization in other parts of speech, see Wegera (1996), Bergmann and Nerius (1998), Rössler (1998), Bergmann (1999), and Szczepaniak and Barteld (2016).

animacy scale. While humans are conceived of as highly individual, which is also reflected in the distinct tendency to use proper names for clear and unambiguous reference, the individual features of each single animal do not play a role in the conceptualization of, say, sheep or fish (see Yamamoto 2008: 44). It has been argued that a higher position on the animacy hierarchy makes an entity more salient and cognitively more accessible (see Zaenen et al. 2004: 118). Thus, the observation that the capitalization of nouns starts out with proper names, i. e., at the highest level of animacy and individuality, and subsequently spreads to words denoting persons, concrete concepts, and abstract concepts (cf. Bergmann and Nerius 1998) fits in well with the hypothesis that capitalization was originally used to highlight particularly salient concepts.

In addition, the data reported on in Bergmann and Nerius (1998) reveal regional differences in the development of sentence-internal capitalization, suggesting a spread from the East to the West of the German-speaking territory (also see Maas 1995b: 91). Both aspects – animacy of referents on the one hand and the area where a text was written on the other – will therefore be taken into account in our study.

An additional, morphological factor in capitalization, which has hardly been considered in previous studies (with the notable exception of Rössler 1998), is compositional complexity. This factor seems highly relevant given the fact that the sixteenth and seventeenth centuries saw a continuous spread of so-called improper compounds, i. e., compounds emerging from syntactic phrases with a preposed genitive attribute such as *[[der Brücke-n] Zoll]* ‘the bridge’s toll’ > *[der [Brücke+n+zoll]]*, lit. ‘the bridge-toll’ (Demske 1999, 2001; Nübling and Szczepaniak 2013). In the sixteenth century, improper compounds showed a strong tendency toward spelling in two words. It was only toward the end of the seventeenth century that their spelling in one word was almost completely established (see Pavlov 1983; Solling 2012). Especially in separately spelled compounds, capitalization can help the reader to decode the morphologically complex unit as one syntactic word. In this respect, Solling (2012: 169–174), drawing on a corpus of printed sermon texts from 1550 to 1710, finds that the capitalization of compounds clearly prevails over non-capitalization throughout the entire period in question. Hence, we will consider compositional complexity as additional factor in our study.

Earlier research has also shown that the spread of capitalization in handwritten texts seems to occur delayed as compared to its spread in printed texts (see Weber 1958: 112; Risse 1980; Bergmann and Nerius 1998: 76). The comparison of capitalization in Luther’s (printed) Bible on the one hand and his (handwritten) letters on the other suggests that the same factors mentioned above might be at play (see Risse 1980; Moulin 1990).

To our knowledge, none of the existing studies has taken lexeme-specific variation into account. However, even a cursory glance at our corpus data reveals that two lexemes with referents at the same level of the animacy scale can vary considerably in spelling even in one text written by the same scribe. Consider, for instance, the spelling of *mann* ‘man’ and *frau* ‘woman’ in Hamm (1592),³ which, judging by the typeface, is written by one single scribe. In the case of *mann*, 9 of 11 occurrences are capitalized, while all 8 occurrences of *frau* are not.

The considerable amount of spelling variation within one lexeme as well as within the same (semantic) group of lexemes suggests that apart from animacy and regional differences, other lexeme-specific factors such as word frequency influence the choice between uppercase and lowercase spelling. Therefore, our corpus study will take these four factors – word frequency, dialectal area, compositional complexity, and animacy – into account and probe for possible interactions between them.

3 The role of frequency in language change

In usage-based approaches, frequency has been established as an important factor determining language change (see, e. g., Diessel 2007). However, the exact influence of frequency is still subject to considerable debate and entails a variety of open questions (see Blumenthal-Dramé 2012). In addition, frequency has been observed to cause different kinds of effects in language change (see Erker and Guy 2012). For instance, Bybee (2003a, 2003b, 2006) argues for high frequency as a major driving force of grammaticalization. At the same time, however, high frequency is said to have a conserving effect, especially so in the case of morphological change where highly frequent forms have been shown to be highly persistent (Bybee 2006). Erker and Guy (2012: 545) observe a third kind of influence:

[F]requency operates by affecting the behavior of the core conditioning factors. Specifically, high frequency either activates or amplifies the predictive power of other constraints. For factors that are activated via interaction with frequency, their effects appear only among high-frequency forms.

In other words, frequency can operate indirectly, i. e., in interaction with other factors. Just as frequency can drive or impede change, we can assume that it can not only activate or amplify the impact of other factors leading to change but

³ We refer to the protocols by the place name and the year of production, e. g., Hamm (1592). See Table 1 for a complete list.

also inhibit this impact. Regarding the spread of capitalization, then, several scenarios are possible: frequency might have a direct (either conserving or causing change) effect on capitalization, it might interact with another factor (again in both directions, i. e., either inhibiting or amplifying this factor), or both. In the case of capitalization, for instance, it is conceivable that a specific factor like animacy affects high-frequency forms to a higher or to a lower degree than low-frequency forms, or vice versa.

In research on the development of sentence-internal capitalization, the effect of frequency has not been addressed so far. However, there have been studies investigating frequency effects in writing (e. g., Sandra et al. 1999; Largy et al. 2007; Kapatsinski 2010). Both Largy et al. (2007) and Kapatsinski (2010) show that there is, unsurprisingly, a higher amount of uncertainty and variation in the spelling of low-frequency items as compared to high-frequency ones, both in children’s writing and in texts written by adults.

Bybee (2006: 727) also argues for frequency-related prototype effects in a category, i. e., “the most frequent member is central and the other members are more marginal”. In the case of capitalization, then, it seems plausible to assume that if a frequent member of a category often occurs capitalized, other members of the category will follow this pattern. Given the observation that nouns referring to animate entities are often capitalized for pragmatic reasons (see the discussion above), and given the observation that the graphemic form of individual lexemes tends to become more stable over the time period under investigation (see Voeste 2008), we can hypothesize that, in the spread of capitalization, highly frequent nouns referring to animate entities (most prototypically, humans) might serve as “attractors” for other nouns.

4 Corpus: Protocols of witch trials from sixteenth and seventeenth century

The present study draws on a corpus of 18 annotated witch trial interrogation protocols written between 1588 and 1630, edited by Macha et al. (2005). The protocols selected for our study are evenly distributed across three time periods (1580–1599, 1600–1619, 1620–1640)⁴ and across six dialectal areas. Each

⁴ However, it should be noted that the distribution of the texts *within* each time period is rather unbalanced at present, as can be seen from Table 1: while the oldest text from the first time period is from 1588 and the newest one from 1600, the last time period currently only covers texts produced between 1628 and 1630.

Table 1: Protocols of witch trials (the core corpus of 18 protocols).

Place	Region	Year	Number of words	Number of scribes	Transcript or copy
Jever	I	1592	1,521	Not specified	Not specified
Meldorf	I	1618	1,002	Not specified	Not specified
Alme	I	1630	1,342	Not specified	Not specified
Perleberg	II	1588	1,326	Town clerk Arnold Crusemark	Not specified
Güstrow	II	1615	1,490	Notarius publicus Nicolaus Wichmann	Not specified
Stralsund	II	1630	1,485	Not specified	Revised transcript
Hamm	III	1592	1,860	Not specified	Not specified
Gaugrehweiler	III	1610	948	Not specified	Transcript
Lemberg	III	1630	1,331	Imperial notary Melchior Wiltperger	Not specified
Georgenthal	IV	1597	1,830	Not specified	Copy
Rosenburg	IV	1618	1,673	Two scribes	Not specified
Ostrau	IV	1628	1,014	Notary and municipal judge Johannes Engelberck	Not specified
Riedlingen	V	1596	1,173	Not specified (one scribe)	Copy
Günzburg	V	1613	1,165	Not specified (one scribe)	Copy
Baden–Baden	V	1628	1,022	Not specified	Not specified
München	VI	1600	1,934	(Probably) two scribes	Not specified
Schweinfurt	VI	1616	1,894	Not specified	Not specified
Bamberg	VI	1628	1,073	Not specified (one scribe)	Not specified

protocol contains between 1,000 and 2,000 words. The regional division is adopted from Macha et al. (2005): North West (Region I), North East (Region II), Middle West (Region III), Middle East (Region IV), South West (Region V), and South East (Region VI). Table 1 gives an overview of the data and provides additional information about the scribes and the situation of production, i. e., whether we are dealing with a transcript, a revised transcript, or a copy of a transcript.

The protocols were semi-automatically tokenized and annotated with sentence boundaries, lemmas, and part-of-speech tags. However, separately spelled compounds like *Teufel-ß dantz* ‘devil’s dance’ (Alme 1630) pose a problem for the tokenization: while compounds are realized as one graphemic word in Modern German (*Teufelstanz*), they are often written as two or more graphemic words in Early New High German (just like, for instance, many English compounds today). In these cases, capitalization does not occur at the first graphemic word. Therefore, our annotation used two different tokenization

variants, one for syntactic and one for graphemic tokens (see Barteld et al. 2014 for details). For the current study, we only use syntactic tokens; accordingly, syntactic tokens like *dantz Leine* ‘dancing line’ (Alme 1630) consisting of two graphemic tokens (*dantz* and *Leine*), in which only the second, but not the first constituent of the compound is capitalized, do not count as word-initial capitalizations.

In addition, the data were manually coded for animacy. There have been many suggestions for the annotation of animacy at different levels of granularity in the corpus-linguistic literature (see, e. g., Zaenen et al. 2004). For this study, a coarse-grained classification with four levels (human, other animates, concrete, and abstract)⁵ was used for two reasons. For one, a too fine-grained categorization scheme may give rise to many questionable categorizations; moreover, variables with too many levels would prove problematic in the model fitting process outlined in the subsequent section given the size of our data.

To be sure, even the encoding of animacy in such a coarse-grained scheme is not problem-free. For example, metonymic items like *Gericht* ‘court’, which may refer to a group of judges or to the place where they pass judgement, could legitimately be categorized both as animate and as inanimate. As the data were annotated in context, the categorization was decided upon on a case-by-case basis by two independent trained annotators. Inter-rater agreement was very high (Cohen’s $\kappa = 0.95$). The 25 instances in which the annotators did not agree have been removed for this study.

As Macha et al. (2005, XXIII) report, it was not possible to clearly differentiate between uppercase and lowercase letters with regard to <v>/<V>, <z>/<Z>, and <h>/<H>. In these cases, lowercase characters are used in the edited protocols (see Szczepaniak and Barteld 2016 for further discussion). Therefore, these characters – altogether 961 tokens – have been excluded from the present study.

For the mixed model reported on below, the remaining sentence-internal common nouns (i. e., syntactic tokens, see the discussion above) were used. Furthermore, proper nouns were excluded as they are almost systematically capitalized. Excluded as well were names of professions if combined with first names as it is not possible to clearly determine whether they denote the profession or a family name. Even in cases like *Caspar wittich, Clauß Gebhardt, so beede Gärtner* ‘Caspar Wittich, Claus Gebhardt and both *Gärtner*’ (Bamberg 1628), they show an ambiguity. Here, it is not possible to clearly determine whether *Gärtner* is a family name or denotes a profession (“gardener”). Interestingly, the names of professions that were retained in our data show a

5 As described below, gender was also included for human (or super-human) entities leading to five categories.

Table 2: Proportion of uppercase tokens for different parts of speech.

PoS	Proportion of uppercase tokens (%)
Common nouns (human)	54
Names of professions	71
Proper nouns	89

higher tendency toward capitalization than other common nouns referring to humans, as Table 2 shows.⁶ This could be interpreted as a graphemic reflection of this ambiguity. The frequent use of lexemes such as *Gärtner* as proper names might enhance their propensity toward capitalization even in cases where they are used as common nouns.

In total, 2,593 sentence-internal common nouns out of 15,535 syntactic tokens (3,045 nouns) were retained in the dataset.

5 Factors influencing capitalization and their interaction

In this section, the individual factors influencing the use of capitals in the protocols of witch trials will be discussed on the basis of our data. Section 5.1 deals with the influence of animacy, Section 5.2 discusses frequency effects and the interaction between animacy and frequency. In Section 5.3, we address variation between different scribes. In Section 5.4, morphological complexity is addressed.

5.1 Animacy as factor influencing graphemic variation

As pointed out in Section 2, previous studies suggest that the spread of sentence-internal capitalization proceeds along a scale of animacy and concreteness: “humans > concrete > abstract”. While it may not be immediately obvious why

⁶ The numbers differ from those reported in Schutzeichel and Szczepaniak (2015), as for the current study we removed all words beginning with <v>, <h>, or <z>, and all instances where the annotators did not agree on animacy. In addition, Table 2 only reports the proportion of uppercase common nouns referring to humans, while Schutzeichel and Szczepaniak (2015) take all common nouns into account.

animacy and concreteness are conflated in one single hierarchy, it seems plausible to assume that both aspects are closely connected. As discussed above, animacy interacts with individuation, i. e., “the degree to which we see something as a clearly delimited and identifiable individual entity” (Yamamoto 2008: 5). As Vogel (1996: 125) points out, animate entities can be seen as “ontologically maximally discrete”. As such, they are prime examples of clearly delimited and identifiable individual entities. Consequently, the hierarchy “animate > concrete > abstract” can, in various guises, be found quite frequently in the relevant literature. Consider, for instance, Langacker’s “empathy hierarchy” (Langacker 1991: 322) given in (1):

- (1) [AN *human* > *animal* AN] > [INAN *physical object* > *abstract entity* INAN]

These distinctions also occur in Zaenen et al.’s (2004) more fine-grained animacy classification scheme. Along with various other categories, we find the hierarchy given in (2).

- (2) HUMAN > ANIMAL > CONCRETE > NONCONC[RETE]

Finally, Ewald (1992) sees living creatures and objects as prototypical cases of concrete entities, without distinguishing between humans and animals. In the classification scheme used by Bergmann and Nerius (1998), by contrast, animals are subsumed under “concrete nouns”, but their distinction “persons > other concrete nouns > abstract nouns” also reflects the hierarchy “animate > concrete > abstract”.

Our data lend further support to the hypothesis that animacy and concreteness play a pivotal role in the development of sentence-internal capitalization. Table 3 gives the percentages of capitalized nouns for animate, concrete, and abstract referents for each of the three periods under investigation.

Table 3: Percentage of uppercase nouns for different semantic categories.

Animate	1580–1605 (%)	1605–1620 (%)	1620–1650 (%)
Animate	52	54	66
Concrete	36	24	37
Abstract	33	28	35

The distribution differs from the results obtained in previous studies (see Section 2) in that we do not observe a monotonic increase of uppercase initials. Instead,

we find a decrease in capitalization of lexemes referring to inanimate entities from the first to the second period. This might be a corpus artefact due to the small amount of data, but it could also suggest that capitalization was used more systematically to distinguish animate from inanimate entities. A Chi-squared test shows that the differences between the distributions are highly significant for each period ($\chi^2 = 35.0$ for the first period, 66.41 for the second, and 72.53 for the third; $df=2$ and $p < 0.01$ for all three periods). The effect sizes are fairly small at first but – in line with our hypothesis – increase with each period (Cramer's $V = 0.18$ for the first period, 0.28 for the second, 0.29 for the third).

Taking the variance between different lexemes at the same level of the animacy scale into account entails a slight shift in perspective. Figure 1 shows the distribution of the relative frequency of sentence-internal capitalization for each lemma attested more than 5 times in the respective period.

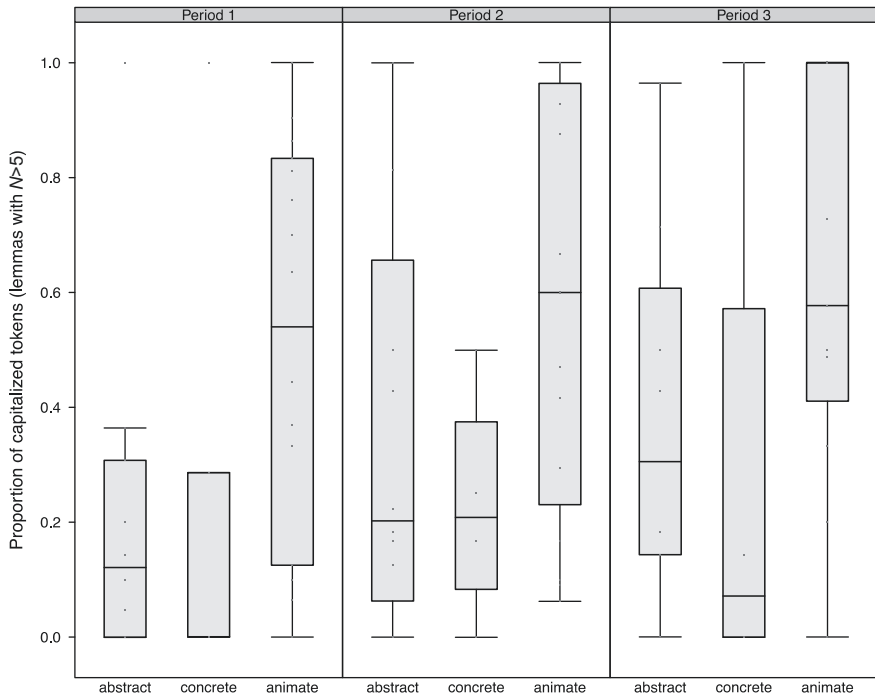


Figure 1: Proportion of capitalized tokens per time frame.

Note that the overall trend, represented by the median, does not indicate any decrease in the capitalization of inanimates from the first to the second period.

In addition, the variance increases, indicating that overall, more lexemes referring to inanimate entities tend toward capitalization. For animate entities, we observe a high variance throughout. However, the distribution concentrates on the upper end in the second and third periods.

A closer look at the data reveals that another factor might play a role in the case of human entities, namely gender. In the case of nouns referring to individual persons, we find a striking difference between nouns denoting men and women. Figure 2 shows the different distributions only for nouns denoting humans, without taking “super-human” beings such as *god* into account. A Fisher’s exact test shows that the difference is highly significant ($p < 0.001$, odds ratio = 3.01; only unambiguous cases considered). If “super-human” beings are taken into account, the distribution remains roughly the same.

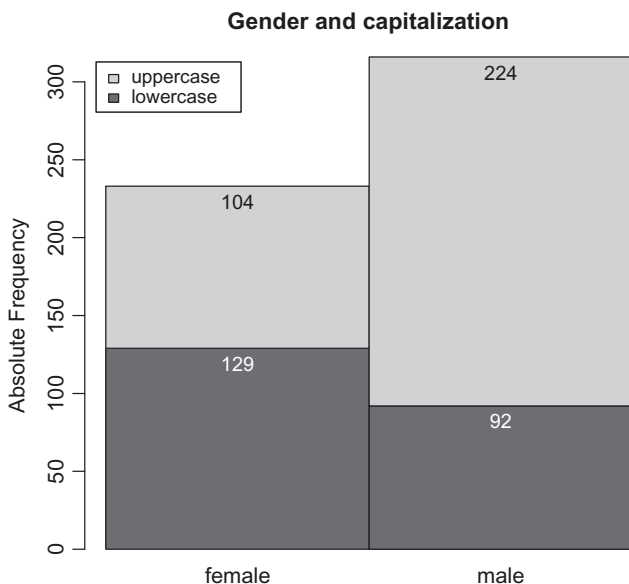


Figure 2: Differences in capitalization for female vs. male referents.

The differences between nouns denoting men and women are hardly surprising not only considering the patriarchal society in which the corpus texts were produced but also the specific characteristics of the text type in question. Across all texts, most of the incriminated persons are women, while the representatives of the legal authorities are men. This obviously amplifies the differences in

social status, showing that, in the texts under investigation, pragmatic factors still play a major role in the choice between uppercase and lowercase initials.

As the gender information is only available for humans, we decided to incorporate it into the variable “Animacy/concreteness”, yielding a five-way hierarchy:

- (3) human or super-human, male > human or super-human, female > animate or human without gender specification (e. g., human collectives) > concrete > abstract.

Admittedly, this way of operationalizing the animacy hierarchy is not entirely unproblematic, especially given that the mid-level category comprises such different entities as animals, human collectives, and individuals without gender specification, e. g., *Kind* ‘child’. However, it can still be considered a good compromise between a too coarse-grained hierarchy on the one hand and a very fine-grained categorization scheme on the other. While the former would lead to many false predictions when used as a fixed effect in a regression model, the latter would lead to a rank-deficient model given the rather small amount of data.

5.2 Frequency effects

When operationalizing lemma frequency, several important choices have to be made. First and foremost, lemma frequency can only be assessed on the basis of a specific corpus. Two options present themselves, then, for the present study: to assess the frequencies of the individual lexemes, we could either use the corpus of witch trial protocols itself, or we could consult another contemporary corpus, ideally a larger one. A larger corpus, such as the *Bonn Early New High German corpus* (<http://www.korpora.org/fnhd/>; FnhdC for short), has the advantage that it might give us a more representative picture of the lemma frequencies. However, using FnhdC would pose several problems. Most importantly, 300,629 of the 515,447 tokens in the corpus are not lemmatized, and 833 of the 852 types attested in our data do not occur in the lemmatized part of the FnhdC corpus at all. In addition, the principles of lemmatization in the FnhdC corpus are quite different from those applied in lemmatizing the SiGS corpus, especially with regard to complex words. For instance, some prefixed words in the FnhdC corpus have been lemmatized with their prefix (e. g., *Begebenheit* ‘happening’ as “be-gebenheit”), others without (e. g., *Beschreibung* ‘description’ as “schreibung”). Therefore, we decided against using the FnhdC frequencies.

As similar problems arise with other publicly available corpora for Early New High German, drawing on the frequencies in the witch trial protocols themselves was the only practicable alternative. This has the additional advantage that these documents might be more representative of conceptually oral language (in the sense of Koch and Oesterreicher 2007) than the data in other available corpora.

5.3 The influence of different scribes

As Bergmann and Nerius (1998) show, areality is another relevant factor for the spread of capitalization in printed texts. Sentence-internal capitalization as an innovation started in the south and spread toward the west and the north of the German-speaking territory. In our corpus, however, there is no clear difference between the areas (Schutzeichel and Szczepaniak 2015). Instead, there are differences between individual protocols. As most of the protocols in the corpus were written by only one scribe, this can be seen as indicating the influence of individual scribes. While printed texts were often corrected by multiple people, handwritten texts show the preferences of individuals. Consequently, we take the influence of individual scribes into account.

5.4 Compositional complexity

One final factor that seems to play a role in our data is the morphological complexity of the individual lexemes. More precisely, compounds are more often capitalized than other lexemes (Figure 3).

Note that the overall number of compounds in the corpus is very small. The difference in the proportion of uppercase vs. lowercase initials between compounds and other lexemes is not very strong but still significant according to a Fisher's exact test ($p < 0.001$); the effect size, however, is very small (odds ratio = 1.96).

While the present data have only been annotated for compounds vs. non-compounds – leaving aside lexicalized and semantically rather opaque items like *Bürgermeister* ‘major’, lit. “citizen-master”, or *Jungfrau* ‘virgin’, lit. “young-woman” – it would be worthwhile to take a closer look at other aspects of morphological complexity in future studies. For example, it would be interesting to investigate whether nouns derived from verbs behave differently from other nouns (see Bergmann and Nerius 1998 for the effect of derivational complexity within the group of abstract nouns).

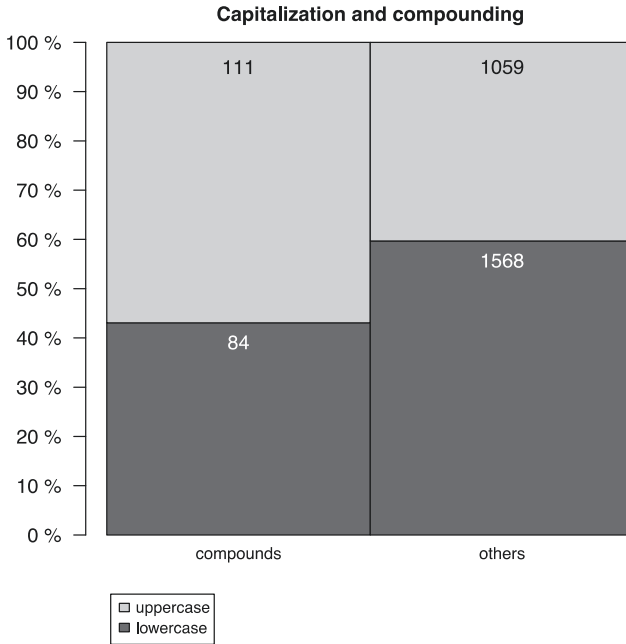


Figure 3: Differences in capitalization for compounds vs. noncompounds.

6 Model fitting and model selection

In the previous sections, we have discussed potential factors playing a role in the choice between uppercase and lowercase initials for common nouns in our data. We use binomial logistic mixed-effects modeling to investigate the influence of the individual variables as well as potential interactions between them.

Mixed-effects models have become increasingly popular in corpus linguistics (Gries 2015: 97). As multilevel models, they allow for taking individual- and group-level variation into account (Gelman and Hill 2007: 246). However, Gries (2015) mentions several open questions with regard to mixed-effects modeling. Most importantly, he points out that the question of how p -values for fixed and random effects are calculated has not been conclusively answered for all models. Furthermore, he indicates that there is no agreed-upon procedure for model selection. In our study, we use R (R Core Team 2015) and *lme4* (Bates et al. 2015) to obtain *Wald* p -values for fixed and random effects, and we use a backward-selection procedure drawing on likelihood-ratio tests to assess the validity of our

models against null models (cf., e.g., Baayen et al. 2008). We start with a maximal model in which FREQUENCY, TIMEPERIOD, LANGUAGEAREA, COMPOSITIONAL COMPLEXITY, and ANIMACY⁷ are entered as fixed effects as well as an interaction term for FREQUENCY and ANIMACY.

In order to take the influence of individual scribes into account, we add PROTOCOL as a random effect into our model. While this is not entirely unproblematic, it provides a good approximation to our target variable, the individual scribe. Unfortunately, scribe changes are not documented in the edition by Macha et al. (2005) for documents written by more than one scribe. Given that only few protocols have been written by more than one scribe, using PROTOCOL as a proxy seems to be an acceptable choice. Discarding all protocols written by more than one scribe would have resulted in considerable data loss, as each combination of period and dialect area is only represented by one text in the corpus. In addition, random slopes were added for the fixed effect of ANIMACY.

Before we turn to the results of the model, let us briefly consider the logic behind choosing the fixed effects and interaction terms just mentioned. As discussed in the previous sections, animacy has been shown to play a role in the emergence of sentence-internal capitalization in printed texts. In addition, we assume that frequency – which has not been considered in previous studies – plays an important role in a writer’s choice between uppercase and lowercase initials. As elaborated on in Section 5.2, frequency can be assumed to either operate on its own or to interact with another factor. For instance, it seems reasonable to hypothesize that frequency either inhibits or enhances the influence of animacy. Even a cursory glance at the data shows that frequency interacts with animacy: among the 10 most frequent lexemes in our corpus, 8 denote animate entities; by contrast, only 72 of 420 hapax legomena refer to animate entities. This is quite unsurprising given that we tend to talk about animate entities, especially about persons, more frequently than about inanimate entities or abstract concepts. Therefore, we added an interaction term FREQUENCY: ANIMACY to the model. Adding random slopes for ANIMACY reflects the idea that the effect of animacy might be different for different scribes. The model specification is given in (4).

$$(4) \quad \text{CAPITALIZATION} \quad \sim \quad \log_{10}(\text{FREQUENCY}) \times \text{ANIMACY} + \text{COMPOUND} + \text{TIMEPERIOD} + \text{LANGUAGEAREA} + (1 + \text{ANIMACY} | \text{PROTOCOL})$$

⁷ Animacy is used as a shorthand for “Animacy/Concreteness” as explained in Section 5.1 above.

Table 4: Results of a binomial mixed-effects model.

	Estimate	Std. Error	z-Value	Pr(> z)	VIF
(Intercept)	-1.37	0.33	-4.19	<0.001***	0
log ₁₀ (Freq)	0.51	0.12	4.32	<0.001***	2.52
Animacy_concrete	0.76	0.21	3.66	<0.001***	4.33
Animacy_animate	1.00	0.33	3.02	<0.001***	6.3
Animacy_human_female	2.93	0.52	5.64	<0.001***	9
Animacy_human_male	1.77	0.43	4.08	<0.001***	5.37
compound_YES	0.66	0.18	3.65	<0.001***	1.1
Period_2	-0.33	0.36	-0.9	0.37	1.35
Period_3	0.22	0.38	0.58	0.56	1.36
LanguageArea_WEST	0.18	0.30	0.6	0.55	1.01
log ₁₀ (Freq):Animacy_conc	-1.11	0.25	-4.5	<0.001***	2.87
log ₁₀ (Freq):Animacy_anim	-0.24	0.25	-0.94	0.35	4.91
log ₁₀ (Freq):Animacy_hum_fem	-1.93	0.32	-6.08	<0.001***	7.42
log ₁₀ (Freq):Animacy_hum_male	-0.04	0.23	-0.19	0.85	2.78

The results of this maximal model are given in Table 4.

Following a model selection procedure along the lines of Baayen et al. (2008), we can arrive at a leaner model by taking out nonsignificant predictors. In the case of our data, `TIMEPERIOD` and `LANGUAGEAREA` do not seem to make a difference. A likelihood ratio test shows that a model without these factors does not perform significantly worse than the maximal model ($\chi^2 = 2.34$, $p = 0.50$). Importantly, it also performs better than a null model without `FREQUENCY` ($\chi^2 = 60.21$, $p < 0.001$) or without the interaction between `FREQUENCY` and `ANIMACY` ($\chi^2 = 56.7$, $p < 0.001$). The index of concordance C , which assesses how well the model predicts the data (cf. Baayen 2008: 204), indicates a better than random fit ($C = 0.76$), which, according to Hosmer and Lameshow (2000: 162), qualifies as “acceptable discrimination”.

All Variance Inflation Factors (VIFs) – which are used to check the assumption that there is no multicollinearity between the predictors – are below the threshold of 10, which is often mentioned as a rule of thumb in the statistical literature (see Levshina 2015: 272). However, it should be mentioned that other approaches set a threshold of 5 and that more conservative approaches advise caution if the average VIF is above 1 (see Field et al. 2012: 276; cf. also O’Brien 2007 for a discussion of such rules of thumb). On the other hand, Levshina (2015: 272) points out that “logistic regression is quite robust with regard to some correlation between predictors.”

In order to check the robustness of the model, a bootstrapping approach was used. In particular, 5,000 items were sampled from the data (with

replacement), and a model was fit to the sample data using the leaner model specification given in (5) resulting from the model selection process outlined above. The procedure was repeated 1,000 times.

$$(5) \quad \text{CAPITALIZATION} \sim \log_{10}(\text{FREQUENCY}) \times \text{ANIMACY} + \text{COMPOUND} + (1 + \text{ANIMACY}|\text{PROTOCOL})$$

Table 5 reports the mean estimates, p -values, and VIFs for the resulting models along with the respective standard deviations.

Table 5: Mean results of binomial mixed-effects model with bootstrapping.

	Estimate	SD (Estimate)	Mean p (Pr(> z))	SD (p)	Mean VIF	SD (VIF)
(Intercept)	-1.33	0.11	<0.001***	<0.001	0	0
\log_{10} (Freq)	0.52	0.10	<0.001***	<0.001	2.51	0.08
Animacy_concrete	0.78	0.15	<0.001***	0.01	5.24	0.31
Animacy_animate	1.03	0.22	0.006**	0.03	8.38	0.68
Animacy_hum_female	3.05	0.40	<0.001***	<0.001	11.6	1.17
Animacy_hum_male	1.84	0.22	<0.001***	<0.001	8.73	1.02
Compound_YES	0.68	0.14	<0.001***	0.005	1.1	0.01
\log_{10} (Freq):	-1.12	0.19	<0.001***	0.003	2.87	0.09
Animacy_concrete						
\log_{10} (Freq):Animacy_anim	-0.24	0.20	0.30	0.29	4.93	0.26
\log_{10} (Freq):	-2.01	0.28	<0.001***	<0.001	7.45	0.93
Animacy_hum_fem						
\log_{10} (Freq):	-0.05	0.18	0.49	0.29	2.67	0.10
Animacy_hum_male						

The bootstrapped model presents most of the trends detected in the model reported in Table 4. Overall, high frequency seems to lead to an increase in capitalization, as does concreteness and even more so animacy. Again, we also find significant interactions between ANIMACY and FREQUENCY. The mean C -value of all bootstrapped models is 0.77 (SD = 0.007), indicating a better than random fit. However, there might still be important factors that have not been included in the model. A closer look at the predictions of the full model outlined above (see Table 4) can therefore prove insightful, as it allows for a more fine-grained comparison of predicted and observed data. Table 6 lists the lexemes with a token frequency of at least 5 that show the highest deviation between estimated

Table 6: Lexemes with the highest deviation between estimated and observed values according to the full model reported in Table 4.

Lemma	Freq.	Uppercase (Obs.)	Uppercase (Pred.)	Error (Abs.)	Error (Rel.)
<i>amt</i> 'office/function'	7	7	0	7	1
<i>kristall</i> 'crystal'	5	5	0	5	1
<i>marter</i> 'torture'	6	6	0	6	1
<i>pulver</i> 'powder'	8	8	0	8	1
<i>gericht</i> 'court'	6	5	0	5	0.83
<i>kläger</i> 'plaintiff'	5	5	1	4	0.8
<i>aussage</i> 'testimony'	15	13	2	11	0.73
<i>ander</i> 'other _N '	7	4	3	5	0.71
<i>pferd</i> 'horse'	24	16	3	17	0.71
<i>erde</i> 'earth/soil'	9	6	0	6	0.67
<i>seele</i> 'soul'	5	3	0	3	0.6
<i>mensch</i> 'human being'	12	8	5	7	0.58
<i>abend</i> 'evening'	7	3	1	4	0.57
<i>garten</i> 'garden'	7	6	2	4	0.57
<i>salbe</i> 'balm'	7	4	0	4	0.57
<i>milch</i> 'milk'	9	5	0	5	0.56
<i>krankheit</i> 'illness'	11	6	0	6	0.55
<i>feind</i> 'enemy'	13	1	13	12	0.92
<i>leute</i> 'people'	21	3	13	12	0.57
<i>katze</i> 'cat'	6	0	3	3	0.5
<i>witwe</i> 'widow'	13	5	6	5	0.38
<i>kind</i> 'child'	30	5	12	11	0.37
<i>morgen</i> 'morning'	6	0	2	2	0.33
<i>kuh</i> 'cow'	29	13	19	8	0.28
<i>stück</i> 'piece'	8	0	2	2	0.25
<i>teufel</i> 'devil'	53	28	34	12	0.23
<i>sohn</i> 'son'	37	28	30	6	0.16
<i>frau</i> 'woman'	68	26	29	11	0.16
<i>bekennnis</i> 'confession'	20	5	6	3	0.15
<i>geist</i> 'ghost/spirit'	34	23	28	5	0.15
<i>gott</i> 'god'	41	37	39	6	0.15
<i>mann</i> 'man'	38	33	36	5	0.13

and observed values.⁸ The first part of the table gives lexemes for which the amount of capitalization is underestimated, the second part gives lexemes for which it is overestimated by the model.

⁸ The error values in Table 6 give the number of tokens where the prediction of the model for or against capitalization, using a cutoff value of 0.5, is wrong. Both absolute and relative frequencies of the errors are given. Compounds have been disregarded in this table.

Most of the lexemes where the model underestimates the degree of capitalization either denote concepts that are central to the frame of court interrogations (cf. *kläger* ‘plaintiff’, *amt* ‘office/ official function’). Or they denote key concepts in the crimes the accused are recurrently charged with, e. g., using a magic balm (*salbe*) or powder (*pulver*), causing or magically healing an illness (*krankheit*), or depriving cows of their capacity to give milk (*milch*). These cases point toward the importance of thematic emphasis, which underscores the finding from previous research that the use of capital letters was largely pragmatically conditioned in its initial stages.

The number of capitalized instances is overestimated in the case of some concrete nouns like *stück* ‘piece’, animates like *katze* ‘cat’, and human or super-human entities like *feind* ‘enemy’, *teufel* ‘devil’, or *geist* ‘ghost’. Some of these examples might point to the importance of taking the context into account as well, rather than just individual words: for example, *feind* ‘enemy’ exclusively occurs in the NP *der böse feind* ‘the evil enemy’, referring to the devil. In one case, the adjective, rather than the noun, is capitalized. In addition, the model’s underperformance in the case of *kind* ‘child’ and *kuh* ‘cow’ might indicate the relevance of semantic roles, especially of agentivity. Children and cows are often the victims of the accused’s alleged witchery (e. g., *habe Sie [...] in des bösen namen ain Khu geschlagen* ‘she had slaughtered a cow in the evil’s [= devil’s] name’, Baden-Baden 1628; *Vnd das Sie [...] Ihr aigen khindt vmbgebracht* ‘and that she killed her own child’, Baden-Baden 1628). Conversely, among the lemmas often wrongly predicted to occur in lowercase, there are many which tend to exhibit a high level of agency in the respective context.

Another factor that has not been taken into account in the present study is the emotional value of the lexemes. Note that many of the lexemes occurring in lowercase more often than expected are connoted negatively. Obviously, *feind* ‘enemy’ entails a highly negative connotation, as does *teufel* ‘devil’. Also, cats (*katze*) are only mentioned in direct connection with the defendants, who have allegedly used them as riding animals (sic!, Jever 1592, Alme 1630) or as ingredients for magic potions (Lemberg 1630). Again, this can be seen as indicating the key role of pragmatic factors in the choice of capital letters in the time period under investigation.

7 Conclusion: Theory and data

In this paper, we have discussed various factors determining the use of sentence-internal capitalization of nouns in Early New High German witch

interrogation protocols. Previous studies, conducted mainly on the basis of printed texts, have shown animacy and areality to be relevant factors. However, the substantial amount of variation between lexemes belonging to the same animacy class has been neglected so far. Therefore, a more fine-grained approach is necessary. Current multivariate and multilevel statistical methods allow for taking different factors and their interaction into account. Using binomial logistic regression, we explored various factors that can plausibly be assumed to play a role along with animacy/concreteness.

In sum, we have been able to identify some additional significant predictors for capitalization. Our study provides further evidence for the role of animacy in the spread of sentence-internal capitalization, with our five-scale animacy hierarchy turning out to be a significant predictor of capitalization. Another significant predictor was compositional complexity. Furthermore, our data point to an effect of frequency. However, frequency closely interacts with other factors, some of which have been captured in our model, while others must be left for future research. The interaction patterns between frequency and animacy suggest that, roughly speaking, cognitive and pragmatic factors enhance or inhibit the degree of capitalization especially in the domain of high-frequency items: both for inanimates and for terms denoting women, capitalization becomes less likely with increasing frequency.

A *post-hoc* inspection of the full model has suggested that some factors not considered in the present model might play a role as well. For instance, terms denoting highly agentive entities or concepts with a high degree of salience in the context of witch interrogations seem to be particularly prone to capitalization, whereas negative emotional value appears to inhibit capitalization. In addition, the (syntactic) context has not been taken into account in the present study (apart from excluding sentence-initial tokens). However, the grammatical function of a word might play a role as well. Note that Langacker (1991) discusses the empathy hierarchy – his rough equivalent to the animacy hierarchy – in connection with the prototypical grammatical subject.

Overall, however, our study shows that the domain of graphemic variation stands to benefit from an integration of multivariate quantitative methods on the one hand and philological diligence on the other. So far, inquiries into the usage and spread of German sentence-internal capitalization have almost exclusively relied on qualitative and monofactorial quantitative analyses. Our study has explored the potential of going beyond monofactorial approaches in assessing diachronic changes in the German writing system. However, it has also become clear that using mixed-effects modeling for corpus data entails certain challenges and open questions. This is especially true for historic texts, where the

amount of available data are small and the creation of reliably annotated corpora is more expensive than for modern texts.

Future research should follow up on these open questions and challenges. First and foremost, however, a larger amount of data would be needed in order to arrive at a more fine-grained picture of the various factors involved in the emergence and spread of sentence-internal capitalization. Given a sufficient amount of diachronic data, the inventory of different methods used to assess these data could be expanded as well. For example, growth curve models and additive modeling could be used for diachronic data in which the temporal dimension does make a difference (which was not the case for our data).

While some questions regarding the usage and spread of capitalization in Early New High German have to be left open, our results – along with the findings of previous studies – demonstrate that graphemic data can play an important role in uncovering the cognitive mechanisms involved in language use and language change. On the one hand, the motivations for capitalization discussed in this paper, i. e., cognitive and pragmatic factors such as animacy, have been identified as major driving forces determining language users' choice between different variants and ultimately leading up to language change elsewhere as well. To mention just one example, Köpcke (2000) has shown animacy to play a key role in the reorganization of the German declension class system. On the other hand, the diachronic spread of sentence-internal capitalization is a prime example of *conventionalization*, which in turn is a key mechanism in the emergence and diachronic change of linguistic signs (see Traugott and Trousdale 2013).

In this regard, studying cases in which historical uses of capitalization deviate from the emerging convention can prove quite insightful. Therefore, future research should expand the scope of investigation from (common) nouns to other parts of speech such as pronouns and adjectives. While these parts of speech are never capitalized in Modern German orthography (with the exception of the honorific pronoun *Sie* 'you' and, in some cases, its less formal counterpart *Du*), they tend to occur quite frequently with uppercase initials in the Early New High German data. Investigating the rise and fall of these capitalization patterns can also shed light on factors driving the capitalization of nouns, which started out as a pragmatic strategy and came to be an orthographic rule.

Resources

The dataset and R script have been made available at the Tromsø Repository for Language and Linguistics (TroLLing): <http://dx.doi.org/10.18710/SJ4OQE>.

References

- Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R. Harald, Doug J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4). 390–412.
- Barteld, Fabian, Renata Szczepaniak & Heike Zinsmeister. 2014. The definition of tokens in relation to words and annotation tasks. In Verena Henrich, Erhard Hinrichs, Daniël de Kok, Petya Osenova & Adam Przepiórkowski (eds.), *Proceedings of the thirteenth International Workshop on Treebanks and Linguistic Theories (TLT13)*, 250–258. Tübingen: Department of Linguistics, University of Tübingen. <http://tlt13.sfs.uni-tuebingen.de/tlt13-proceedings.pdf>
- Bates, Douglas, Martin Mächler, Benjamin Bolker & Stephen Walker. 2015. *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1–9. <https://CRAN.R-project.org/package=lme4>.
- Bergmann, Rolf. 1999. Zur Herausbildung der deutschen Substantivgroßschreibung. Ergebnisse des Bamberg-Rostocker Projekt. In Walter Hoffmann, Jürgen Macha, Klaus J. Mattheier, Hans-Joachim Solms & Klaus-Peter Wegera (eds.), *Das Frühneuhochdeutsche als sprachgeschichtliche Epoche: Werner Besch zum 70. Geburtstag*, 59–79. Frankfurt am Main: Peter Lang.
- Bergmann, Rolf & Dieter Nerius. 1998. *Die Entwicklung der Großschreibung im Deutschen von 1500–1700*, 2 vols. Heidelberg: Winter.
- Blumenthal-Dramé, Alice. 2012. *Entrenchment in usage-based theories: What corpus data do and do not reveal about the mind*. Berlin: De Gruyter Mouton.
- Bock, Michael, Klaus Hagenschneider & Alfred Schweer. 1989. Zur Funktion der Groß- und Kleinschreibung beim Lesen deutscher, englischer und niederländischer Texte. In Peter Eisenberg & Hartmut Günther (eds.), *Schriftsystem und orthographie*, 23–56. Berlin: De Gruyter.
- Buringh, Eltjo & Jan Luiten Van Zanden. 2009. Charting the ‘Rise of the West’: Manuscripts and printed books in Europe, A long-term perspective from the sixth through eighteenth centuries. *The Journal of Economic History* 69(2). 409–445.
- Bybee, Joan. 2003a. Cognitive processes in grammaticalization. In Michael Tomasello (ed.), *The new psychology of language*, Vol. II, 145–167. New Jersey: Lawrence Erlbaum Associates Inc.
- Bybee, Joan. 2003b. Mechanisms of change in grammaticization: The role of frequency. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 602–623. Oxford: Blackwell.
- Bybee, Joan. 2006. From usage to grammar: The mind’s response to repetition. *Language* 82(4). 711–733.
- Dahl, Östen. 2008. Animacy and egophoricity: Grammar, ontology and phylogeny. *Lingua* 118(2). 141–150.
- Dahl, Östen & Kari Fraurud. 1996. Animacy in grammar and discourse. In Thorstein Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility*, 47–65: John Benjamins.
- Demske, Ulrike. 1999. Case compounds in the history of German. In Matthias Butt & Nanna Fuhrhop (eds.), *Variation und Stabilität in der Wortstruktur: Untersuchungen zu*

- Entwicklung, Erwerb und Varietäten des Deutschen und anderer Sprachen*, 150–176. Hildesheim: Olms.
- Demske, Ulrike. 2001. *Merkmale und Relationen: Diachrone Studie zur Nominalphrase im Deutschen*. Berlin: Mouton de Gruyter.
- Diessel, Holger. 2007. Frequency effects in language acquisition, language use, and diachronic change. *New ideas in psychology* 25(2). 108–127.
- Dietrich, Rainer & Kathy Y. van Nice. n. d. Belebtheit, Agentivität und inkrementelle Satzproduktion. Unpublished manuscript.
- Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik. Band 1: Das Wort*. 4th updated and revised edn. Stuttgart: Metzler.
- Erker, Daniel & Gregory R. Guy. 2012. The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language* 88(3). 526–557.
- Ewald, Petra. 1992. Konkreta versus Abstrakta: Zur Semantischen Subklassifikation deutscher Substantive. *Sprachwissenschaft* 17. 259–281.
- Fuhrhop, Nanna. 2009. *Orthografie*. 3rd updated edn. Heidelberg: Winter.
- Gauger, Hans-Martin. 1994. Geschichte des Lesens. In Otto Ludwig & Hartmut Günther (eds.), *Schrift und Schriftlichkeit: An interdisciplinary handbook of international research*, Vol. 1, 65–84. Berlin: De Gruyter.
- Gelman, Andrew & Jennifer Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
- Gfroerer, Stefan, Hartmut Günther & Michael Bock. 1989. Augenbewegungen und Substantivgroßschreibung – Eine Pilotstudie. In Peter Eisenberg & Hartmut Günther (eds.), *Schriftsystem und orthographie*, 111–136. Berlin: De Gruyter.
- Gries, Stefan Th. 2015. The most under-used statistical method in corpus linguistics: Multi-level (and mixed-effects) models. *Corpora* 10(1). 95–125.
- Hosmer, David W. & Stanley Lemeshow. 2000. *Applied logistic regression*, 2nd edn. New York: Wiley.
- Field, Andy, Jeremy Miles & Zoë Field. 2012. *Discovering statistics using R*. Los Angeles: Sage.
- Kapatsinski, Vsevolod. 2010. What is it I am writing? Lexical frequency effects in spelling Russian prefixes: Uncertainty and competition in an apparently regular system. *Corpus Linguistics and Linguistic Theory* 6(2). 157–215.
- Kaempfert, Manfred. 1980. Motive der Substantiv-Großschreibung: Beobachtungen an Drucken des 16. Jahrhunderts. *Zeitschrift für Deutsche Philologie* 99. 72–98.
- Knoop, Ulrich. 1994. Entwicklung von Literalität und Alphabetisierung in Deutschland. In Otto Ludwig & Hartmut Günther (eds.), *Schrift und Schriftlichkeit: An Interdisciplinary Handbook of International Research*, Vol. 1, 859–872. Berlin: De Gruyter.
- Koch, Peter & Wulf Oesterreicher. 2007. Schriftlichkeit und kommunikative Distanz. *Zeitschrift für germanistische Linguistik* 35(3). 346–375.
- Köpcke, Klaus-Michael. 2000. Starkes, Schwaches und Gemischtes in der Substantivflexion des Deutschen: Was weiß der Sprecher über die Deklinationsparadigmen? In Rolf Thieroff, Matthias Tamrat, Nanna Fuhrhop & Oliver Teuber (eds.), *Deutsche Grammatik in Theorie und Praxis*, 155–170. Tübingen: Niemeyer.
- Labs-Ehler, Brigitte. 1993. *Versalschreibung in althochdeutschen Sprachdenkmälern: Ein Beitrag über die Anfänge der Großschreibung im Deutschen unter Berücksichtigung der Schriftgeschichte*. Göppingen: Kümmerle.
- Langacker, Ronald W. 1991. *Foundations of cognitive grammar. Vol. 2: Descriptive application*. Stanford: Stanford University Press.

- Largy, Pierre, Marie-Paule Cousin, Peter Bryant & Michel Fayol. 2007. When memorized instances compete with rules: The case of number-noun agreement in written French. *Journal of Child Language* 34(2). 425–437.
- Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.
- Maas, Utz. 1992. *Grundzüge der deutschen orthographie*. Tübingen: Niemeyer.
- Maas, Utz. 1995a. Einige Grundannahmen zur Analyse der Groß- und Kleinschreibung im Deutschen, insbesondere zu ihrer Grammatikalisierung in der Frühen Neuzeit. In Gotthard Lerchner, Marianne Schröder & Ulla Fix (eds.), *Chronologische, areale und situative Varietäten des Deutschen in der Sprachhistoriographie: Festschrift Für Rudolf Große*, 85–100. Frankfurt am Main: Peter Lang.
- Maas, Utz. 1995b. Bäuerliches Schreiben in der Frühen Neuzeit: Die Chronik des Hartich Sierk aus den Dithmarschen aus der ersten Hälfte des 17. Jhds. In Wolfgang Raible (ed.), *Kulturelle Perspektiven auf Schrift und Schreibprozesse*, 65–96. Tübingen: Narr.
- Macha, Jürgen, Elvira Topalović, Iris Hille, Uta Nolting & Anja Wilke (eds.). 2005. *Deutsche Kanzleisprache in Hexenverhörprotokollen der frühen Neuzeit, 2 Vols*. Berlin: De Gruyter.
- Moulin, Claudine. 1990. *Der Majuskelgebrauch in Luthers Deutschen Briefen: (1517–1546)*. Heidelberg: Winter.
- Munske, Horst Haider. 1995a. Überlegungen zur Rechtschreibreform und zur Rekonstruktion der wortbezogenen Groß- und Kleinschreibung.: Ein Diskussionsbeitrag zu Utz Maas' Aufsatz "Rechtschreibung und Rechtschreibreform". *Zeitschrift für Germanistische Linguistik* 23(1). 59–68.
- Munske, Horst Haider. 1995b. Zur Verteidigung der deutschen Orthographie: Die Groß- und Kleinschreibung. *Sprachwissenschaft* 20. 278–322.
- Nübling, Damaris & Renata Szczepaniak. 2013. Linking elements in German: Origin, change, functionalization. *Morphology* 23. 67–89.
- O'Brien, Robert M. 2007. A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity* 41(5). 673–690.
- Pavlov, Vladimir M. 1983. *Zur Ausbildung der Norm der deutschen Literatursprache im Bereich der Wortbildung (1470–1730): Von der Wortgruppe zur substantivischen Zusammensetzung*. Berlin: Akademie-Verlag.
- R Core Team. 2015. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Risse, Ursula. 1980. *Untersuchungen zum Gebrauch der Majuskel in deutschsprachigen Bibeln des 16. Jahrhunderts: Ein historischer Beitrag zur Diskussion um die Substantivgroßschreibung*. Heidelberg: Winter.
- Rössler, Paul. 1998. Die Großschreibung in Wiener Drucken des 17. und frühen 18. Jahrhunderts. In Werner Bauer & Hermann Scheuringer (eds.), *Beharrsamkeit und Wandel: Festschrift für Herbert Tatzreiter zum 60. Geburtstag*, 205–238. Wien: Edition Praesens.
- Sandra, Dominiek, Steven Frisson & Frans Daems. 1999. Why simple verb forms can be so difficult to spell: The influence of homophone frequency and distance in Dutch. *Brain and Language* 68. 277–283.
- Sattler, Johann Rudolph. 1975 [1617]. *Teutsche Orthographe und Phraseologie*. Reprint of the edition Basel, 1617. Hildesheim: Olms.
- Schutzichel, Marc & Renata Szczepaniak. 2015. Die Durchsetzung der satzinternen Großschreibung in Norddeutschland am Beispiel der Hexenverhörprotokolle. In Markus

- Hundt & Alexander Lasch (eds.), *Deutsch im Norden: Varietäten des norddeutschen Raumes im Spiegel der germanistischen Sprachgeschichtsschreibung*, 151–167. Berlin: De Gruyter.
- Solling, Daniel. 2012. *Zur Getrennt-, Zusammen- und Bindestrichschreibung von Substantivkomposita im Deutschen (1550–1710)*. Uppsala: Acta Universitatis Upsaliensis.
- Szczepaniak, Renata. 2011. Gemeinsame Entwicklungspfade im Spracherwerb und im Sprachwandel? Kognitive Grundlagen der onto- und historiogenetischen Entwicklung der satzinternen Großschreibung. In Klaus-Michael Köpcke & Arne Ziegler (eds.), *Grammatik – Lehren, Lernen, Verstehen*, 341–360. Berlin: De Gruyter.
- Szczepaniak, Renata & Fabian Barteld. 2016. Hexenverhörprotokolle als sprachhistorisches Korpus. In Sarah Kwekkeboom & Sandra Waldenberger (eds.), *PerspektivWechsel oder: Die Wiederentdeckung der Philologie. Vol. 1: Sprachdaten und Grundlagenforschung in Historischer Linguistik*, 43–70. Berlin: Erich Schmidt.
- Topalović, Elvira. 2003. *Sprachwahl – Textsorte – Dialogstruktur: Zu Verhörprotokollen aus Hexenprozessen des 17. Jahrhunderts*. Trier: WVT Wissenschaftlicher Verlag.
- Voeste, Anja. 2008. *Orthographie und Innovation: Die Segmentierung des Wortes im 16. Jahrhundert*. Hildesheim: Olms Verlag.
- Traugott, Elizabeth C. & Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.
- Vogel, Petra M. 1996. *Wortarten und Wortartenwechsel: Zur Konversion und verwandten Erscheinungen im Deutschen und in anderen Sprachen*. Berlin: De Gruyter.
- Weber, Walter Rudolf. 1958. *Das Aufkommen der Substantivgroßschreibung im Deutschen: Ein historisch-kritischer Versuch*. München: Uni-Druck.
- Wegera, Klaus-Peter. 1996. Zur Geschichte der Adjektivgroßschreibung im Deutschen: Entwicklung und Motive. *Zeitschrift für Deutsche Philologie* 115. 382–392.
- Yamamoto, Mutsumi. 2008. *Animacy and reference: A cognitive approach to corpus linguistics*. Amsterdam: John Benjamins.
- Zaenen, Annie, Jean Carletta, Gregory Garretson, Joan Bresnan, Andrew Koontz-Garboden, Tatiana Nikitina, M. Catherine O'Connor & Tom Wasow. 2004. Animacy encoding in English: Why and how. In Bonnie Webber & Donna Byron (eds.), *DiscAnnotation '04*, 118–125. Stroudsburg, PA: Association for Computational Linguistics.