



Historical  
Sociolinguistics and  
Authorship Elucidation  
in Medieval Private  
Written Correspondence:  
Theoretical and  
Methodological  
Implications for and  
from Forensic  
Linguistics\*

**JUAN A. CUTILLAS-ESPINOSA &  
JUAN M. HERNÁNDEZ-CAMPOY**

---

\* Financial support for this research has been crucially provided by the DGICT of the Spanish Ministerio de Economía y Competitividad (FFI2014-56084-P) and by Fundación Séneca (19331-PHCS-14), the Murcian Agency for Science and Technology (Programas de Apoyo a la Investigación).

**Abstract** Corpora of historical private and official correspondence and their substantial social metadata from the past offers a very useful archival source to carry out studies in Historical Sociolinguistics. However, illiteracy among female population and the subsequent use of scribes in remote periods make authorship and gender constitute some of the most controversial socio-demographic issues when doing sociohistorical research. Letters might not have been autographs but rather written by a scribe by way of dictation, which can lead to the distortion of findings concerning authorship and gender-based patterns, from the perspective of sociolinguistic variation. On the other hand, Forensic Linguistics appeared as a branch of Applied Linguistics to assist the law in legal processes, where authorship elucidation is often one of the most disputed questions. In this paper we will present an overview of the main approaches to authorship attribution within Forensic Linguistics and relate them to sociohistorical data in the case of the letters by Margery Paston, putting their theoretical tenets and techniques to the test of time. The data suggests that formal (spelling) features are less indicative of authorship than other morphosyntactic markers. Forensic Linguistics and Historical Sociolinguistics can mutually benefit each other, by sharing their expertise in authorship research and its application to current and historical texts in their social context

**Keywords:** forensic linguistics, historical sociolinguistics, authorship, holograph/autograph letters, scribal/secretarial letters

### **1. Introduction: Authorship in Historical Private Correspondence**

Authorship constitutes one of the most controversial issues when doing socio-historical research on the behaviour of linguistic forms. The possibility of retrieving personal information in private correspondence has facilitated the reconstruction of some of the socio-demographic variables traditionally correlated with linguistic production, provided that the surviving letters are autographed documents and their written words were the proper ‘utterance’ of their authors, with no mediation of scribes or secretaries. This is an important methodological aspect with crucial theoretical and particularly

phenomenological consequences. As we will see below in section 3.1, letters often were not autographed but rather written from dictation by a personal scribe of the author due to the widespread illiteracy that was characteristic of early historical periods (see Hernández-Campoy and Conde-Silvestre 1999; Schneider 2002; Bergs 2005; 2015; Nobels & van der Wal 2009; Hernández-Campoy 2016b; or van Hattum 2017).

Solving the mystery of the presence of authors' or the scribes' personal language use and practices in non-autographed letters can just be an exercise of historical reconstruction of sociolinguistic scenarios with fragmentary data from the long distance in time tunnel, and always under the risk of anachronisms due to the possible distortion of an inadequate application of the uniformitarian principle to the speech community and also the communities of practice. With this in mind, the aim of this paper is twofold. First of all, we aim to present an example of historical epistolary documents where forensic linguistic techniques may help to elucidate authorship: letters authored by Margery Paston. Secondly, we will try to show how this set of texts offers a valuable perspective on the different approaches to forensic authorship identification mentioned before. We are, therefore, both applying forensic linguistic techniques to historical data and evaluating the results with the aim of providing insight for Historical Sociolinguistics and its written materials. A brief summary of the main theoretical tenets of *Historical Sociolinguistics* and *Forensic Linguistics* will be presented first as a theoretical framework.

## **2. Theoretical Framework: Historical Sociolinguistics and Forensic Linguistics**

### **2.1. Historical Sociolinguistics and Private Correspondence**

#### **2.1.1. Historical Sociolinguistics**

*Historical Sociolinguistics* appeared in the early 1980s as the confluence of Historical Linguistics and Sociolinguistics with the aim of applying the tenets and findings of contemporary sociolinguistic research to the interpretation

of linguistic material from the past. This practice has allowed scholars the reconstruction of patterns of language variation and change longitudinally in earlier speech communities. Retrospectively, works such as Romaine (1982), Milroy (1992), Nevalainen & Raumolin-Brunberg (1996, 1998, 2003/2017, 2012), Ammon, Mattheier & Nelde (1999), Jahr (1999), Kastovsky & Mettinger (2000), Bergs (2005), Hernández-Campoy & Conde-Silvestre (2012), Russi (2016) or Säily, Nurmi, Palander-Collin & Auer (2017) reflect the vitality of this impressively burgeoning field.

However, representativeness and statistical validity have often been questioned in Historical Sociolinguistics, making Labov (1994: 11) assert that this discipline constitutes “the art of making the best use of bad data” (see also Labov 1972; Nevalainen & Raumolin-Brunberg 2003/2017, 2012; or Hernández-Campoy & Schilling 2012). The methodological difficulties are mostly due to the fact that the sociolinguistic study of historical language forms inevitably has to rely on the only available linguistic records from previous periods – most of which will be incomplete, fragmentary, or non-representative in some way – as well as on socio-historical (and cultural) backgrounds that can only be reconstructed rather than directly observed or experienced by the researcher (see also Raumolin-Brunberg 1996; Nevalainen & Raumolin-Brunberg 1998, 2003; Nevalainen 1999; Ayres-Bennett 2001; Schneider 2002; or Bauer 2002). In addition to these problems, the *historical paradox* and the so-called *uniformitarian* principle have also been part of the controversial issues in its methodological procedure (see Bergs 2012 and Conde-Silvestre & Hernández-Campoy 2012). Admittedly, it is true that this is an exercise of socio-historical reconstruction of the remote past time of a given language through the use of materials whose size is inevitably limited for different reasons (Schneider 2002: 89). This means, as Bergs (2005: 71) points out, that the non-existence of evidence does not obviously allow for empirically irrefutable conclusions about the non-existence of individual facts.

These obstacles in the historical sociolinguistic practice have largely been overcome thanks to the assistance of *Corpus Linguistics*, *Computational Linguistics* and *Social History*, which have conferred empirical reliability upon the discipline (see Nevalainen & Raumolin-Brunberg 2003/2017; Hernández-Campoy & Conde-Silvestre 2012; or Säily, Nurmi, Palander-Collin & Auer 2017). Advancements in computing technology as well as the

development of digitised resources and large text corpora have been radically transforming linguistic research since the 1980s (see Cantos & Sánchez 2000; Cantos 2012). The compilation of large electronic corpora (both computer-driven and research-driven) – containing collections of naturally occurring language data – has also been highly instrumental in overcoming some of the problems inherent to working with ‘bad data’ from the past and the historical paradox (Sebba & Fligelstone 2001; Bauer 2002; Schneider 2002; Bergs 2007b; Baker 2010; Murphy 2010; or Cantos 2012, 2013).

### 2.1.2. Historical Corpora of Private Correspondence

The development and diversification of archival data sources is allowing scholars to explore the role of new genres and text-types as adequate materials for sociolinguistic analysis: ego-documents, such as diaries, travel accounts, court records, recipes, and especially letters, are now seen as essential documents for research in this field at diastatic, diatopic and diaphasic levels (see Elspaß 2002, 2012; Tieken-Boon van Ostade 2005, 2006; Nevala & Palander-Collin 2005; Palander-Collin, Nevala & Nurmi 2009; Palander-Collin 2010; Auer 2015; Schiegg 2016; Krogull, Rutten & van der Wal 2017; Voeste 2018; or Hernández-Campoy & García-Vidal 2018a, 2018b, among others). Some monographs have also confirmed the relevance of these documents to reconstruct the sociolinguistic contexts of language variation and change in the past (see Dossena & Fitzmaurice 2006; Nevalainen & Tanskanen 2007; Dossena & Tieken-Boon van Ostade 2008; Sairio 2009, 2017; Dossena & Del Lungo Camiciotti 2012; van der Wal & Rutten 2013; Rutten & van der Wal 2014; Auer, Schreier & Watts 2015).

Unlike other surviving documents from the past, collections of both private and official correspondence, have favoured the interest in linguistic variation in the history of languages given the different socio-demographic and geographical characteristics of writers and recipients. In fact, private letters from historical corpora constitute the language production that is closest to Labov’s (1966) everyday speech, so their study may shed light onto the resources and driving forces for sociolinguistic variability in remote societies (Nevala & Palander-Collin 2005; Nevalainen & Tanskanen 2007; Palander-Collin 2010; Conde-Silvestre & Hernández-Campoy 2013). The

study of historical letters has meant a crucial contribution to the detection of the social provenance and direction of long-term changes longitudinally and macroscopically (see Biber 1995: 283–300; 2001: 98–99; Biber & Finegan 1989, 1997; Nevala & Palander-Collin 2005; Nevalainen & Tanskanen 2007; Palander-Collin 2010; Conde-Silvestre & Hernández-Campoy 2013), as well as to the analysis and reconstruction of the sociolinguistic behaviour of individual speakers in social interaction microscopically (Palander-Collin 1999; Palander-Collin, Nevala & Nurmi 2009; Auer 2015; Hernández-Campoy & Conde-Silvestre 2015; Conde-Silvestre 2016; or Hernández-Campoy & García-Vidal 2018), consolidating the historical validity of some ‘sociolinguistic universals’ (see Nevalainen & Raumolin-Brunberg 1996; 2003/2017; Hernández-Campoy & Conde-Silvestre 2012).

However, an important problem related to the authenticity and purity in the transmission of manuscripts is authorship and, hence, gender. As seen above, letters were often not autographed but rather written from dictation by a personal scribe of the author due to the widespread illiteracy that was characteristic of early historical periods. Illiteracy among female population and the subsequent use of scribes make authorship and gender constitute some of the most controversial socio-demographic issues when doing sociohistorical research, since it might easily corrupt the representativeness and validity of any empirical research carried out ignoring this usual practice.

## 2.2. Forensic Linguistics and authorship identification

Forensic Linguistics appeared as a branch of Applied Linguistics to assist the law in legal processes. The birth of Forensic Linguistics as a discipline is linked to the pioneering work of Jan Svartvik (1968) and his analysis of authorship attribution. Thus, the question of who wrote what has been a key issue in forensic linguistic research since its beginnings – as in Shakespearean authorship disputes (Hope 1994) in the 16<sup>th</sup> century or the possible multi-authorship in the Pentateuch according to the German priest H. B. Winter in the 18<sup>th</sup> century (Olsson 2004: 11). This interest in the discipline has also been reflected in the publication of introductory books and handbooks devoted to this subject (see Coulthard & Johnson 2007, 2010; Olsson 2004, 2008; Olsson & Luchjenbroers 2014; or Turell 2005, among others).

But determining authorship has been regarded as a particularly controversial area within Forensic Linguistics. While admitting that it is possible to identify a particular speaker/writer by their linguistic choices, forensic linguists have disagreed about what particular methodology should be applied. The main assumption about the possibility of determining authorship by linguistic means is the existence of an *idiolect*. This idiolect has to be understood as the set of personal linguistic choices, which, in a sense, constitute the author's signature. The idiolect is certainly not as unequivocal as a DNA sample but, in spite of that, it remains a valuable source of information. The ways in which this idiolect may be characterized vary greatly and, simplifying somewhat, we have divided the approaches to authorship attribution into two main groups. The first one, Forensic Stylistics, is data-driven, in the sense that it identifies salient features of a text that have not been established beforehand. Forensic Stylistics does not normally do statistics. The second one, Forensic Stylometrics, uses pre-selected features (very often morphosyntactic in nature) and does statistics, which is one of its greatest strengths.

### 2.2.1. Characterizing the Speaker's Idiolect

Assuming the concept of idiolect, most forensic linguists claim that, in one way or another, every speaker has their own individual sociolinguistic behaviour that manifests itself in language production through distinctive and idiosyncratic choices (see Coulthard 2004: 431–432). It is commonplace to compare authorship identification techniques with other forensic sciences, such as DNA or fingerprint analysis. In this way, a set of pre-selected linguistic features are aprioristically selected to be observed in a given text (see Turell 2010 or Wright 2013, for example). However, as Coulthard (1994) points out, the 'idiolect' has not been unanimously accepted as an identifying factor, since forensic linguists are unsure whether it is possible to distinguish one person's language from that of a comparable member of his speech community. Identifying a disputed author through his/her idiolect is by no means an easy task. As Coulthard suggests, trying to compare the task of forensic linguists with that of fingerprint analysts is actually counterproductive. It would be difficult (if not impossible) to establish a corpus of texts so big that it would enable us to successfully compare the style of the suspect with

that of the general population. Fortunately, though, forensic linguists are rarely confronted with such a daunting task. Normally, the forensic expert “is not being asked to positively identify an author, but rather to provide grounds to *doubt* that a given suspect was the author. Even when a positive identification is required, other evidence has always massively reduced the list of possible authors, sometimes to only two [...], and rarely to more than a dozen” (Coulthard 1994: 31).

Turell (2010) expanded this idea by suggesting the concept of an ‘idiolect style’ defined as ‘the set of options that writers take from the linguistic repertoire available to them as users of a specific language’ (Turell 2010: 217). The identification of authorship markers starts by choosing the relevant ones for a particular case and then submitting them to empirical evaluation through statistical methods. These markers may also convey sociolinguistic information about the author. When texts are very short, it may prove necessary to combine both qualitative and quantitative methods (like Discriminant Function Analysis), as well as to use data from corpora (Coulthard 1994) to determine the relative rarity of a particular linguistic element. Wright (2013), however, studies the concept of idiolect based on the greetings and farewells of a corpus of e-mails from the Enron company. Less common choices proved to be diagnostic of authorship, even when compared to a corpus of 126 e-mails. All this research seems to suggest that the idea of an idiolect may be useful for authorship identification. The exact nature of the speaker’s choices, though, is open to discussion: some researchers have looked at texts searching for relevant markers, even if these were not amenable to statistical analysis; others have insisted on establishing markers a priori and providing statistical information.

### 2.2.2. Forensic Stylistics

Forensic stylistics was first proposed in work by McMenemy (1993, 2001, 2002, 2004, 2010). He claimed that any writing sample exhibits the author’s personal style with their individual patterns of sociolinguistic variation and aggregate sets of habitual linguistic choices taken from the stock of linguistic alternatives held in common within the speech community (McMenemy 2010: 492). Forensic stylistics makes use of the observation of variables



such as i) spelling; ii) punctuation; iii) non-standard grammatical features (use of *who* instead of *whom*; use of *which* as a relative pronoun when the antecedent is a person; non-native use of verb tenses); iv) variation in morphology (presence or absence of appropriate *-ed* inflections; presence or absence of plural inflection; plural for possessive or vice versa; confusion of adjective and adverbial forms); v) variability in syntactic structure (e.g. order in ditransitive verb constructions; lack of subject/verb agreement, multiple levels of embedding), or vi) discourse markers. Forensic Stylistics is data-driven and recognises features of the particular piece of evidence studied. That is, rather than pre-selecting particular markers to be analysed, “the style markers used for analysis of a particular set of writings must be first observed as linguistic variables in those very writings” (McMenamin 2010: 505), proceeding in a bottom-up fashion. Forensic Stylistics looks at the text first and then decide what salient features may be potential style markers (see also Howald 2008).

The greatest criticism of Forensic Stylistics comes from its inability to provide an error rate, that is to say, statistical support that may be presented to court as evidence. This is particularly relevant in the case of the United States legal system, where evidence presented to court has to comply with the so-called Daubert criteria (see Tiersma & Solan 2002). There have been attempts, though, to apply statistical methods to forensic stylistic studies, thus reconciling two apparently contradictory approaches to authorship (Grant 2007; and Nini & Grant 2013).

### 2.2.3. Forensic Stylometrics

Purely statistical analyses of texts can be traced back to Augustus de Morgan’s suggestion in the 19<sup>th</sup> century that different biblical authors could be identified by word length. Thus, mean word length could be used as an indicator of authorship, as Mosteller & Wallace’s (1964) and Morton & Michaelson (1990) suggested (see also Grant & Baker 2001; Grant 2007, 2010; or Coulthard & Johnson 2007). Chaski (2001, 2005, 2007) is one of the supporters of these stylometric approaches that select a priori markers of authorship and then submits them to statistical analysis. Her methodology proposed is based on the quantitative analysis of the use of syntactically classified punctuation

and syntactic analysis of phrase structure. According to Chaski (2001), these are aspects of linguistic performance unlikely to be consciously modified and, consequently, potentially good markers of authorship – unlike the tests of sentence complexity, vocabulary richness and content analysis used in Forensic Stylistics, which may be unsuccessful in short texts and incapable of meeting the Daubert criteria.

McMenamin (2001: 93) criticised “Chaski’s limited historical perspective and her narrow linguistic focus”, as well as her use of a very limited corpus (four writers) to try to identify markers that are valid for all speakers. Grant & Baker (2001) respond to Chaski’s views by criticising the validity of her method based on syntax only. They claim that “it is not enough to simply show that a particular marker works or does not work in a particular case, as it does not follow that the marker will or will not work in *all* cases” (Grant & Baker 2001: 77). Instead, they defend the application of Principal Component Analysis (PCA) to identify which potential combination of markers could determine authorship in a particular case: “given a set of texts by two unknown authors and a query text, known to be by one of the authors, PCA creates groupings (representing authors) according to markers or combinations of markers, which can indicate which author the query text belongs to” (Grant & Baker 2001: 68). This PCA approach has been successfully applied by Rico-Sulayes (2011) to a set of texts obtained from online forums about drug trafficking in Mexico and by Zhang (2016) to the attribution of a set of Chinese short e-mails combining pragmatic, discourse semantic and discourse information.

Apart from providing other relevant information for historical sociolinguists, our paper will also contribute to determine which one of the two approaches presented before could be more appropriate for authorship attribution of texts from the past.

### **3. Evidence from the *The Paston Letters*: Authorship in Margery’s Letters**

The *Paston Letters* is a collection of 422 authored documents written by 15 members belonging to different generations of this family mainly during

the fifteenth century (from 1425 to 1503), with roughly 246,353 words<sup>1</sup>. The Pastons is the most well documented gentry family of late medieval England that takes its name from a Norfolk village about 20 miles north of Norwich (see Davis 1954; 1971; Richmond 1990/2002, 1996; Castor 2004; Bergs 2005, or Wood 2007). Since Blume's (1882) study on the language of the *Paston Letters*, the corpus has been subject to a variety of studies from different linguistic and stylistic stances (Davis 1954, 1967; Schäfer 1996; Tanabe 1999; Wood 2007; among others), as well as through sociolinguistic approaches (Davis 1983; Bergs 2005; Hernández-Campoy & Conde-Silvestre 1999; Conde-Silvestre & Hernández-Campoy 2004; Hernández-Campoy 2008, 2013, or Hernández-Campoy & Vidal-García 2018).

Many of the letters were autographed, but others were dictated and written by a scribe. According to Davis (1954, 1965, 1971) and Cressy (1980), Clement II, John II, John III, Walter and William III, for example, wrote their own letters themselves, but other family members, such as Agnes, Margaret, Elizabeth or Margery used scribes due to their illiteracy (see Bergs 2005: 79).

### 3.1. Literacy and Dictation in Medieval England

Literacy attainment in the late Middle Ages and the Early Modern period of England was inextricably linked to social position – a complex mixture of wealth and status – as well as to social conventions and locality (see Power 1922/1964; Baldwin 1943, 1944; Stone 1964, 1969; Duckett 1965; Simon 1966; Orme 1973, 1984, 2006; Cressy 1980, 1981, 1993; O'Day 1982; Shahar 1983; Crawford 1985; Houston 1985; Vincent 1989; Davis & Joyce 1989; Archer 1992; Goldberg 1992; Barratt 1992/2010; Cherewatuk & Wiethaus 1993; O'Mara 1996; Wheale 1999; Finke 1999; Daybell 2001; or Krug 2002). Conventionally, literacy was linked to gender, since women's subordination by patriarchal hierarchy meant a serious barrier to their access to education and literacy: conventional beliefs and attitudes of the time assumed that women

---

<sup>1</sup> The body of letters used for the present study was taken from the Middle English Collection of the Internet electronic edition of the *Paston Letters* (First Part) from the Virginia University Electronic Text Center (<http://etext.lib.virginia.edu/toc/modeng/public/PasLett.html>) and also from the Corpus of Middle English Prose and Verse at the University of Michigan (<http://quod.lib.umich.edu/c/cme/paston>).

were subservient to men, with their labour being focused on household management and childraising, and thus not having opportunities for education – at least at the level of functional literacy (see Power 1922: 260–84, 568–81; Orme 1973: 54–55; or Barratt 1992/2010: 2). Before 1500 most English society was illiterate (Bergs 2015: 115): 10% of male population and only 1% of females could sign their own name (Cressy 1980: 177). At the beginning of Elizabeth's reign (1558–1603), only 20% English males and 5% females had signature literacy (Barratt 1992/2010: 261). In the early 17<sup>th</sup> century literacy increased to 33% in men and 10% in women, whereas by the 18<sup>th</sup> century it was about 50% and 20% respectively (see also Cressy 1981, 1993; Vincent 1989:1; or Wheale 1999: 43).

From the point of view of social status, literacy for the nobility and aristocracy was a functional skill and a competence but it was also used non-functionally as a social marker: though male literacy was functional and complete (both reading and writing), in women it was mostly for the cultivation of reading. Children belonging to these elite groups of noblemen were taught by private tuition at home or in another noble household, usually by the family chaplain, to whatever level their families found as necessary (see Orme 1984: 156–160), but only boys were prepared for university at fashionable schools such as Shrewsbury or Westminster (see Wheale 1999: 50).

This problem in authorship and the social category of gender would definitely affect the representativeness of informants and thus the reliability of results from a socio-demographic perspective. For this reason, great care has to be taken in interpreting gender-based patterns of variation, especially those found in the writings of female informants, given that there is sociohistorical evidence strongly suggesting that female informants did not usually write the letters themselves in those periods, but the family clerk, chaplain, or other scribes connected to the family. Bergs (2005; 2015) has dealt with forensic linguistic aspects in late medieval times, and so have Nobels & van der Wal (2009) and van Hattum (2017) in the 16<sup>th</sup>–17<sup>th</sup> and 19<sup>th</sup> centuries respectively. Aware of the problems with authorship and gender in epistolary corpora, Nobels & van der Wal (2009) developed the Leiden Identification Procedure (LIP) in order to differentiate autographical from non-autographical letters, which enabled them to study the sociolinguistic behaviour of authors accurately and reliably as well as the encoding practices

of both social and professional scribes. Van Hattum (2017) has recently applied Forensic linguistics to the interpretation of threatening notices in 19<sup>th</sup> century rural Ireland through Critical Discourse Analysis.

As some studies have demonstrated through quantitative analysis (see Bergs 2005, 2015; or Hernández-Campoy 2016b), it seems that verbatim dictation would reflect the morphosyntactic or lexical variables of the person dictating, but not the relevant phonological or graphological forms. This would crucially help in the forensic elucidation of authorship in medieval private written correspondence, since grammatical and lexical forms would be part of the usual author's sociolinguistic practices (and competence), whereas orthography would belong to the scribe's graphological habits. Bergs (2005; 2015) studied the use of the *h-/th-* personal pronouns and the relative pronoun *which* by the different members of the Paston family in their letters. During the Middle English period, the OE southern Anglo-Saxon *h*-pronouns underwent a process of suppletion in the 3<sup>rd</sup> person plural of all genders, so that those *h*-forms were gradually substituted by the OE northern Scandinavian *th*-pronouns, which were derived from Old Norse. The historical OE *h*-forms had completely been replaced by Scandinavian *th*-forms in the late 15<sup>th</sup> century, which were eventually incorporated into the incipient standard English (see Moore & Marckwardt 1981: 94; Lass 1992: 120; Bergs 2005: 83–103; Brinton & Arnovick 2006/2011: 288–290, among many others). In the case of the relative pronoun *which*, it was commonly used with both animate and inanimate antecedents during the 14<sup>th</sup> century. But the interrogative *who* acquired the function of a relative connector with animate antecedents in the following century, restricting the use of *which* to only inanimate ones (see Bergs 2005: 133–144; also Pyles & Algeo 1964; Millward 1989; Lass 1992, 1994, 2000; Blake 1992, 1996; or Brinton & Arnovick 2006/2011, for example). In his results, as a 15<sup>th</sup> century Norfolk family and, dialectologically, East Midland speakers, the Pastons produced both *h-* and *th-*pronoun forms: 39% were the conservative *h*-forms and 61% were the innovating *th*-forms. In the results from his study of animacy of the form *which* in relative clauses, the Pastons also exhibited presence of both the conservative and the innovating variants with ANimate and INAnimate antecedents: 23% were with animate and 77% with inanimate antecedents.

Longitudinally, he found a communal change by means of the successive addition of generational shifts, with a clear pattern of intermediate stages

of presence of both forms between the categorical use of the conservative variant and that of the innovative one at the other end (Bergs 2005: 103). Cross-sectionally, despite this generational tendency, inter- and even intra-generational differences revealed a much less uniform picture, full of variation, with innovators, fast adopters, and also lames. In fact, quite a different picture emerges when we zoom in on individual members of the family considering the question about authorship and scribal practices. According to Wood (2007: 50), at least 29 different hands made up the letter collection attributed to Margaret Paston, for example. In her study, using a critical discourse analysis approach, she suggests that “even with the formulaic openings of the letters there is evidence that Margaret was partly responsible for the wording” (Wood 2007: 51). If we accept, as Bergs (2005: 128) does, based on Davis (1971, 1954) and Cressy 1980), that Clement II acted as scribe to Agnes’s dictated letters, and Edmond, John II and John III to their mother (Margaret), in the case of the use of the innovative 3<sup>rd</sup> person plural TH-pronoun, their frequencies do not show similar tendencies, which substantiates the hypothesis that different scribes have only little influence on the morphosyntax and lexical shape of the letters: Agnes Paston exhibits 33% of the innovative TH-pronoun in her dictated letters while her amanuensis (Clement II) has only 17% in his autographed letters; Margaret obtains a frequency of 45% whereas her scribes use the same variant 97% (John II), 100% (Edmond II), and 69% (John III). Similarly, in the case of the use INAnimate *which*, there are also dramatic divergent patterns of sociolinguistic behaviour between Margaret (19%) and his sons John II (78%), Edmond II (68%), and John III (75%). Therefore, the influence of these female Pastons as authors of their correspondence on their amanuenses is unequivocal. The most illuminating individual example is the case of Edmond II and usage of the 3<sup>rd</sup> person pronoun variable, since he does not show any traditional *h*-forms but only *th*-pronouns (100%), and, contrarily, when he is taking down Margaret’s dictation, he makes use 45% of the same innovating Scandinavian variant. This means that when Edmond II is writing his own letters, he categorically uses the *th*-forms 100% of the time, but when he is writing for his mother, the presence of these forms is dramatically reduced to 45% and that of the conservative *h*-pronouns increases from 0% up to 55% (Bergs 2005: 124–128; 2015).

### 3.2. The Case of Margery Paston

Margery Brews (?1455–1495) was the daughter of Sir Thomas Brews of Topcroft in Norfolk, and married John Paston III in 1477. She lived in Norwich first and later in London. Six letters are preserved under her name and they are all addressed to John III, two when still unmarried and four after marriage. According to Davis (1954, 1965, 1971) and Cressy (1980), her letters were probably not written by the sender (non-autographed). Although Margery used to dictate her letters, she was the only female family member who was, at least, able to sign her own name (Davis 1971: xxxvii; Bergs 2005: 79).

#### 3.2.1. A Stylometric Approach to Orthographic and Grammatical Forms

For the linguistic analysis of Margery's letters, a pre-determined set of linguistic features (research-driven), implying some statistics, will be followed, as practiced in stylometric approaches. The variables used here as 'distinguishing traits' for authorship attribution in Margery's letters are both orthographic and grammatical forms (see Table 1):

**Table 1: Variables used for text analysis in Margery Paston's Letters.**

Level	Variable	Forms	Status	Examples
Orthography	þ	<þ>	non-standard	<i>þerfor, broþer, thynkeþ</i>
		<th>	standard	<i>therfor, brother, thynketh</i>
	3	<3>	non-standard	<i>3e, 3ou, 3ow, 3oure(s), 3owre(s)</i>
		<y>, <j>, <g>	standard	<i>ye, you, yow, youre(s), yowre(s)</i>
Grammar	2 <sup>nd</sup> p. Y-forms	th-pronouns	non-standard	<i>thou, thow, þou, þow, thee, þee</i>
		y-pronouns	standard	<i>you, yow, ye</i>
	3 <sup>rd</sup> p.pl. TH-pronouns	h-pronouns	non-standard	<i>he(o), ho, ha, hi, hy, hjo/h3o, hir(e), heore, har, her(e), heres, heren, heom, him/hem, hom, ham</i>
		th-pronouns	standard	<i>thai, they, thair, thaim, them, tham(e)</i>
	Which	ANimate antecedent	non-standard	<i>... Syr Richard Egecum wech londyd in Breten ...</i>
		INanimate antecedent	standard	<i>... of youre welfar, the wheche I beseche Allemyghty God preserue ...</i>

The orthographic variables used are the ‘thorn’ (þ) and the ‘yogh’ (ȝ). During the Old English period, the Celts and the Anglo-Saxons used an alphabet of Runes. But the early Christian missionaries introduced the Roman alphabet when they brought Christianity, literacy and European culture to England during the early 7<sup>th</sup> century A.D. (see Upward & Davidson 2011; Crystal 2012). The adoption of the Roman alphabet at the expense of the Runic one was rapid except for a few letters that did not have an equivalent in Latin and thus were still prevailing until the end of the Middle Ages: ‘wynn’ ƿ (>’uu/w’), ‘eth’ ð (>’th’), ‘yogh’ ȝ (‘y/j/g’), and ‘thorn’ þ (>’th’). In the case of the the old runic spelling <þ>, the process of replacement by the digraph <th> taken from the Roman alphabet was completed in the late 15<sup>th</sup> century. As found in Stenroos (2004, 2006), Bergs (2007a), Jensen (2012) and Conde-Silvestre & Hernández-Campoy (2013), the presence of <th> in both Latin and Biblical texts acted as an influential external prestigious norm, so that the Roman-based orthographic form became overtly popular during the 15<sup>th</sup> century as a historical change operating above the level of social awareness and in connection with social and stylistic factors. The use of <th> was gradually supralocalised into the literate ranks of the whole of England: its sociolinguistic diffusion initially took place in the careful and conscious styles, acquiring overt prestige and becoming part of the accepted linguistic norm, as a typical Labovian ‘change from above’. Therefore, the use of (TH) was a sociolinguistic variable with status of *marker* and indexical meaning in late medieval England (Hernández-Campoy & García-Vidal 2018a, 2018b). Similarly, although use of the ‘yogh’ ȝ had some more complex phonotactic constraints, it was substituted with the graphemes <y>, <j> or <g> (see Scragg 1974: 10; Benskin 1977: 506–507; 1982: 18–19; Stenroos 2006; Bergs 2007a; or Conde-Silvestre & Hernández-Campoy 2013, among others).

The grammatical variables used are the 2<sup>nd</sup> person pronominal forms (Y), 3<sup>rd</sup> person plural pronominal forms (TH) and the relative pronoun which. The *y*-based plural forms *you/yow/ȝou/ȝow* and *ye/ȝe* for the 2<sup>nd</sup> person pronouns began to be used with a singular reference in the latter half of the 13<sup>th</sup> century under the influence of the French distinction *vous/tu* at the expense of the *th*-forms (*thou/thow/þou/þow*, *thee/þee*). In the case of variable (TH), during the Middle English period, the OE southern Anglo-Saxon *h*-pronouns underwent a process of suppletion in the 3<sup>rd</sup> person plural of all genders: *h*-forms were

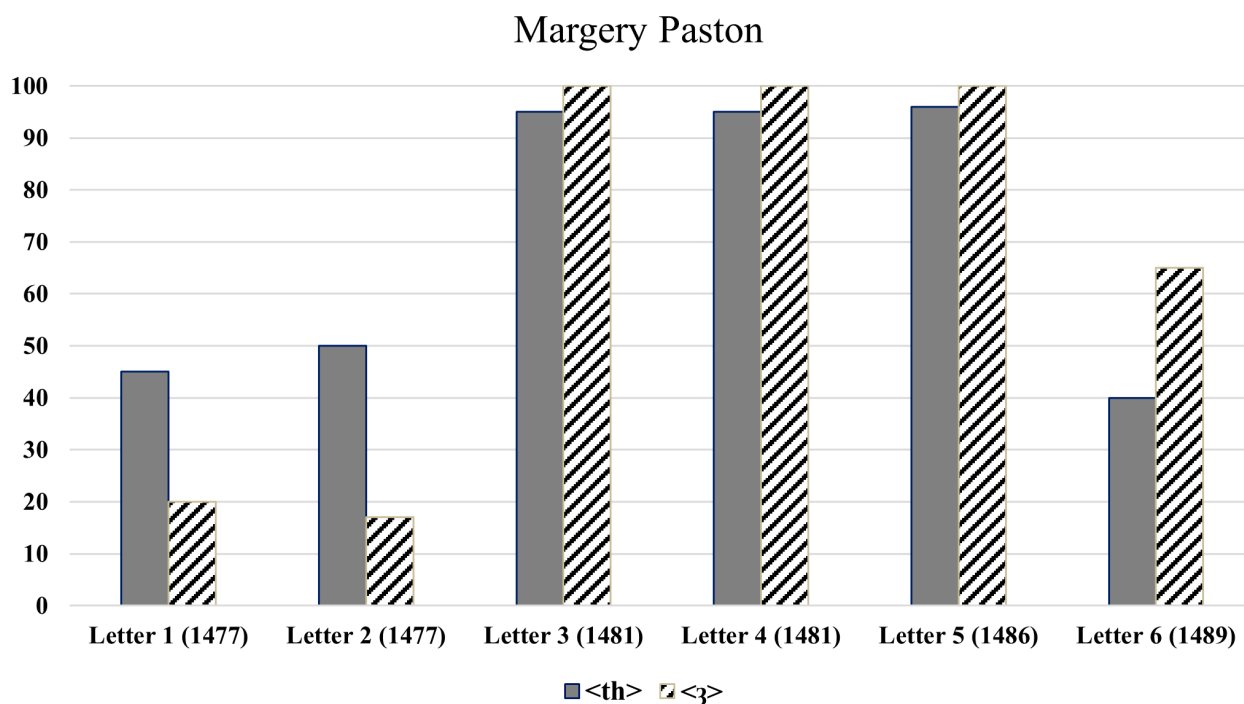


gradually substituted by the OE northern Scandinavian *th*-pronouns (both with *þ* initially and *th* spelling later), which were derived from Old Norse (Lass 1992: 120; Bergs 2005: 83–103; among others). *H*-pronouns were used in the south of England as autochthonous Anglo-Saxon forms (*he* in the South-West, *hēo/ho* in the West Midlands, or *hi/hy* in South-East, for example). Conversely, the Scandinavian *th*-pronouns were salient features in the North and East Midlands, but gradually progressing southwards throughout the Middle English period in a process that took more than three hundred years. Both *h*- and *th*-forms were found in the Paston family correspondence. Finally, the form *which* was used after a preposition, with an antecedent clause in late ME and in non-restrictive relative clauses, taking both animate and inanimate antecedents throughout the period; *whom* and *whos* were commonly used after a preposition and, generally, in non-restrictive relative clauses, also taking animate antecedents; and *what* was used when the antecedent was an indefinite pronoun or a clause. During the 15<sup>th</sup> century, when the interrogative pronoun *who* acquired the function of a relative connector with animate (human) antecedents, then *which* restricted its use to only inanimate (non-human) ones (see Lass 1992; or Bergs 2005: 133–144, for example).

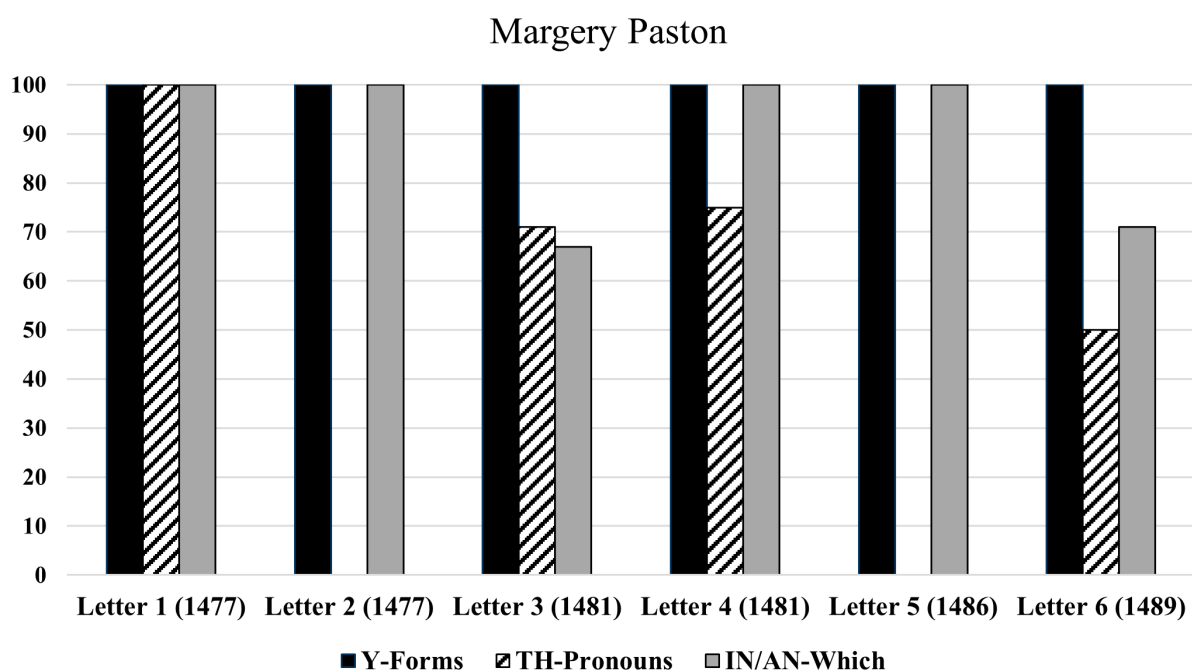
Longitudinally, and macroscopically, in the context of the ongoing communal changes, the examination of the sociolinguistic behaviour exhibited by Margery's life-span (1477–1489) in Table 2 and Figures 1–2 shows inconsistent (even chaotic) patterns, without any kind of developmental linearity or tendency, neither positive nor negative.

**Table 2: Distinguishing traits for Authorship Attribution in Margery's letters.**

Letters	Period	Age	Variant	Orthographic Variables		Grammatical Variables			Totals
				þ	ʒ	2 <sup>nd</sup> p. Y-forms	3 <sup>rd</sup> p.pl. TH-pronouns	IN/AN Which	
<b>Letter 1:</b> #417 (306 words)	1477	22	N-St	55% (16/29)	80% (16/20)	0% (0/16)	0% (0/2)	0% (0/2)	46% (80/148)
			St	45% (13/29)	20% (4/20)	100% (16/16)	100% (2/2)	100% (2/2)	
<b>Letter 2:</b> #418 (339 words)	1477	22	N-St	50% (17/34)	83% (19/23)	0% (0/19)	–	0% (0/3)	54% (54/148)
			St	50% (17/34)	17% (4/23)	100% (19/19)	–	100% (3/3)	
<b>Letter 3:</b> #419 (481 words)	1481	26	N-St	5% (2/44)	0% (0/22)	0% (0/3)	29% (2/7)	33% (1/3)	5% (9/167)
			St	95% (42/44)	100% (22/22)	100% (3/3)	71% (5/7)	67% (2/3)	
<b>Letter 4:</b> #420 (528 words)	1481	26	N-St	5% (3/61)	0% (0/19)	0% (0/3)	25% (1/4)	0% (0/1)	95% (158/167)
			St	95% (58/61)	100% (19/19)	100% (3/3)	75% (3/4)	100% (1/1)	
<b>Letter 5:</b> #421 (221 words)	1486	31	N-St	4% (1/25)	0% (0/22)	0% (0/4)	–	0% (0/1)	2% (1/52)
			St	96% (24/25)	100% (22/22)	100% (4/4)	–	100% (1/1)	
<b>Letter 6:</b> #422 (745 words)	1489	34	N-St	60% (76/127)	35% (8/23)	0% (0/8)	50% (2/4)	29% (2/7)	52% (88/169)
			St	40% (51/127)	65% (15/23)	100% (8/8)	50% (2/4)	71% (5/7)	
<b>Total</b>			N-St	36% (115/320)	33% (43/129)	0% (10/36)	29% (5/17)	18% (3/17)	31% (166/536)
			St	64% (205/320)	67% (86/129)	100% (36/36)	71% (12/17)	82% (14/17)	



**Figure 1: Results of the usage of the orthographic variables in Margery Paston: percentages of the incipient forms per letter.**



**Figure 2: Results of the usage of the grammatical variables in Margery Paston: percentages of the incipient standard forms per letter.**

Averagely, according to aggregate numbers, this age-based pattern might perfectly be a case of age-grading, even though her age-based changes in her sociolinguistic patterns occur after the critical period<sup>2</sup>. Unlike lifespan changes, age-grading relies on individuals changing their speech patterns throughout their lives in a regular and repeated way in each generation: consistent fluctuations in usage are repeated by each generation at a certain age following a chronolectal cyclicity while the community remains stable over time. The Pearson correlation coefficient (Cantos 2013: 58–63) of her scores in a life-context span indicates a positive correlation that might be understood as showing a polynomial relationship along the age-based pattern (with a  $\rho^2$  very close to 1:  $\rho^2=0.9934$ ). This would mean an inverted U or V-shaped curve suggesting age-grading, where the use of the prestige form peaks between her 26 and 31 age, as found here with 95% and 98% use of the standard variants as opposed to 54% in 1477 (age of 22) and 48% (age of 34). The non-prestigious form is used in her early 20s and resurfaced later in her mid 30s (see Figure 3). Inferential statistics through a non-parametric Pearson's Chi-square test of significance (Cantos 2013:75–80) also confirms that the different patterned chronolectal practices did not occurred by chance, with significant relationship at  $p<0.01$  ( $\chi^2=122.363$ ;  $df=3$ ), although data is not robust (often with  $N\leq 5$ ).

---

2 Age, ageing (as both social and biological phenomenon), and hence lifespan (with its three major periods: childhood/adolescence, middle age and old age) have been crucial in the study of variation and change in human language. Pre-adolescence was assumed as the critical period when the brain starts losing plasticity, making language acquisition and learning difficult and much less successful, and thus restraining the likelihood of change in linguistic habits over an adult's life time. However, post-adolescent individuals have been demonstrated to undergo malleability in their sociolinguistic behaviour through life experience and maturation in middle and later stages of adulthood, challenging the conventional assumption in the variationist tradition of uniformity over the lifespan, in addition to pre-adolescence.



**Figure 3: Progression in Margery's sociolinguistic behaviour from 1477 to 1489 in the variables observed.**

Nevertheless, this apparently robust statistical-based diagnosis will prove to be unsuccessful, largely due to the fact that the analysis is carried out in aggregate data. This would question Chaski's (2001) purely quantitative approach. In this case, sociolinguistic information and sociohistorical information may be more crucial for authorship elucidation than mere statistical analysis. There is enough evidence to reject the age-grading hypothesis here. Firstly, there is no gradient progression in the adoption of the standard forms, with sociolinguistic practices drastically increasing from 54% to 95–98% and then decreasing to 45% in a 12-years cohort. Secondly, at the level of individual variables, the patterns of prestige as extralinguistic driving forces do not correlate with the adoption of features, which, otherwise, seems to be chaotic and illogical (especially in the case of grammatical variables). Thirdly, in her 40 years of life-span, it is not likely that Margery would go across only 5 years of age-grading experience (from 1481 at the age of 26 to 1486 at the age of 34) when she was reaching maturity and becoming middle

aged – this phenomenon normally takes some decades in individuals (see Chambers & Trudgill 1980: 91–94). And last, but not least, socio-historical data tells us that Margery Paston was illiterate (see Davis 1971: xxxvii; Cressy 1980; or Bergs 2005: 79). Therefore, these unpredictable patterns with disordered sociolinguistic variation exhibited in Margery Paston from 1477 to 1489 can only be explained by considering the presence of multi-authorship, as an analysis of letters individually and at a micro-level will confirm.

Microscopically, and considering Margery's epistolary interaction and letter recipients (audienceship), her six letters were written when she was between 22 and 34 years old and they were all addressed to the same person (her husband: John Paston III). This means that there was no possibility for distinctions in language production or choice due to addressees' differences in terms of Bell's (1984) audience design or other stylistic resources (see Biber & Finegan 1989; Hernández-Campoy 2016a), as, however, found in other members of the family (see Hernández-Campoy & García-Vidal 2018a, 2018b). Similarly, the short 12 year-span between her first (1477, when she was 22) and last letter preserved (1489, when she was 34, reaching middle age) cannot extend any kind of analysis to chronolectal variation and hence does not allow us to trace any tendencies and propensities in terms of age-grading, for example, as seen above, at the level of individual variables due to its unequal patterning. In fact, inferential statistics on individual variables through Pearson's Chi-square test of significance confirms that Margery's sociolinguistic practices in her letters did not occur by chance in the case of the orthographic variables: the relationship is significant at  $p < 0.01$  ( $\chi^2 = 94.509$ ;  $df = 5$ ) in (b) and at  $p < 0.01$  ( $\chi^2 = 76.252$ ;  $df = 5$ ) in (3). However, in the case of the grammatical variables, their relationship is not significant ( $p > 0.05$ ) and with a small number of tokens (often  $N \leq 5$ ). Moreover, the Pearson correlation coefficient obtained for the relationship among the variables individually in each letter does not indicate a significant correlation ( $p > 0.05$ :  $\rho = 0.1517$ ;  $\rho = 0.0835$ ;  $\rho = 3188$ ;  $\rho = 0.8718$ ; and  $\rho = 0.6288$ ), which means that there is not a function here governed by a predictive model following an implicational scale (the use of variable X goes with the use of variable Y), except in Letter 6 ( $p < 0.01$ :  $r = 0.9833$ ).

Moreover, if macroscopically, at a longitudinal level of language change, the sociolinguistic behaviour was too inconsistent to be indicative of just one

author involved, microscopically, certain overt patterns of sociolinguistic practices can be clearly distinguished. As Figure 1 shows, the production of both orthographic variables (*b* and *3*) in Margery's six letters would confirm that they were not written by herself, but rather by different scribes, whose influence is clearly reflected in the spelling practices. The apparently illogical frequencies obtained for these two variables suggest that there were, at least, two different amanuenses involved: the scribe for letters 1–2 and 6 seems to be different from that for letters 3–5. However, the grammatical variables used in this scrutiny do not shed so much light onto the elucidation and attribution of authorship, as the frequencies of their forms are mostly consistent cross-sectionally, as Table 2 and Figure 2 show. Letters 3–4 and 6 show some presence of the non-standard grammatical 3<sup>rd</sup> person plural pronoun, and letters 3 and 6 also have some cases of the non-standard grammatical WHICH relative. But these results do not exhibit a regular pattern in terms of common practices that could be correlated with spelling use. In grammar, therefore, the influence of the author (dictator) becomes clearly visible on their scribes' hand during dictation – they must be regarded as the authors' and not the amanuenses' personal language use. But quite to the contrary, in no significant case does the orthographic language production of our female informant run counter to their scribes' other usage tendencies – they must be regarded as the amanuenses' and not the authors' personal language use. This fact would confirm, as Bergs (2005: 79) suggested, that scribes were more responsible for the phonological or graphological variables present in verbatim dictated letters, rather than for morphosyntactic and/or lexical features. Assuming that her letters were written by scribes, whose influence would just affect graphological variables, this fact confirms that Margery's letters were written in dictation, and that there were at least two different amanuenses doing it for her.

### 3.2.2. Forensic Stylistic Approach

An approach following Forensic Stylistics and its use of data-driven variables would support the spelling patterns discussed above and the idea of, at least, two hands in Margery's epistolary documents: one of the few linguistic consistencies is, for example, the spelling of the grammatical forms *you* and

*your*, which appear spelt as *yow(e)*, *3e* or *3ow(e)* and *yowre* or *3owr(e)* mostly in Letters 1–2 and 6, whereas as *you*, *ye* and *your* in Letters 3–5 (see Table 4). That is, the most archaic forms (runic symbols) are categorically used in Letters 1–2 and 6, whereas the innovative ones are mostly found in Letters 3–5, which would also suggest that the scribe of letters 1–2 and 6 was much older and dialectal than the scribe of Letters 3–5 – more aware of the incipient national standard English variety.

**Table 4: spelling for 2<sup>nd</sup> person pronoun and adjective.**

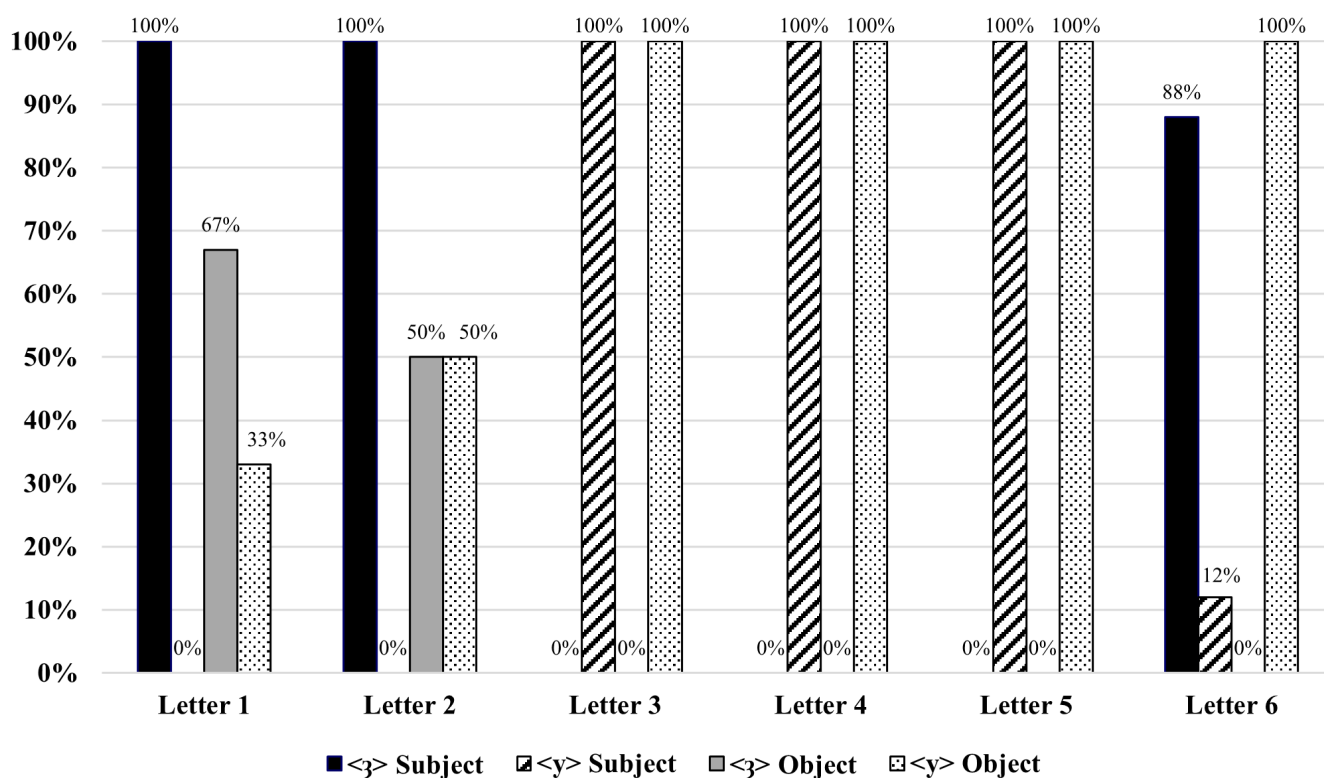
	Pronoun					Adjective			
	<i>you</i>	<i>yow(e)</i>	<i>3ow(e)</i>	<i>ye</i>	<i>3e</i>	<i>your(e)</i>	<i>yowr(e)</i>	<i>3owr(e)</i>	<i>3our(e)</i>
<b>Letter 1</b>	–	3	6	–	7	–	1	1	2
<b>Letter 2</b>	–	4	4	–	11	–	–	3	1
<b>Letter 3</b>	10	3	–	3	–	5	1	–	–
<b>Letter 4</b>	8	–	–	3	–	7	1	–	–
<b>Letter 5</b>	8	–	–	4	–	10	–	–	–
<b>Letter 6</b>	–	8	–	1	7	–	6	–	–

These letters still show the *you-ye* distinction. Historically the second person plural personal pronouns had been the forms *you* for object and *ye* for subject. However, this functional distinction disappeared and *you* replaced *ye* for both functions in late 17<sup>th</sup> century Standard English, after an intermediate period where both forms were indistinctively used for subject and object function (see Nevalainen & Raumolin-Brunberg 2003: 60–61). This grammatical innovation is not relevant here. There is no sign at all of this innovation – even in an embryonic stage –, as all Margery’s letters still exhibit the conventional distinction between object (*you*) and subject (*ye*) forms (see Figure 4).

In terms of forensic authorship theory, while there may be some linguistic markers that are deeply rooted in the speaker’s mind and less likely to be influenced by external factors (dictation), these findings suggest that morphological, syntactic and semantic markers may not exclusively prove



to be more important than purely formal markers (spelling, for example), as Chaski (2001) suggested. The combination of both can be convenient and even technically ideal sometimes. That is, the use of an orthographic form (yogh, for example) may not correlate with the presence of a given grammatical change (*you-ye*, for example), as Figure 4 shows. But the spelling form employed may be determinant when there are no signs of the grammatical innovation yet for author attribution.



**Figure 4: Use of runic symbol yogh <3> and Roman-based <y> in object and subject 2nd person pronouns.**

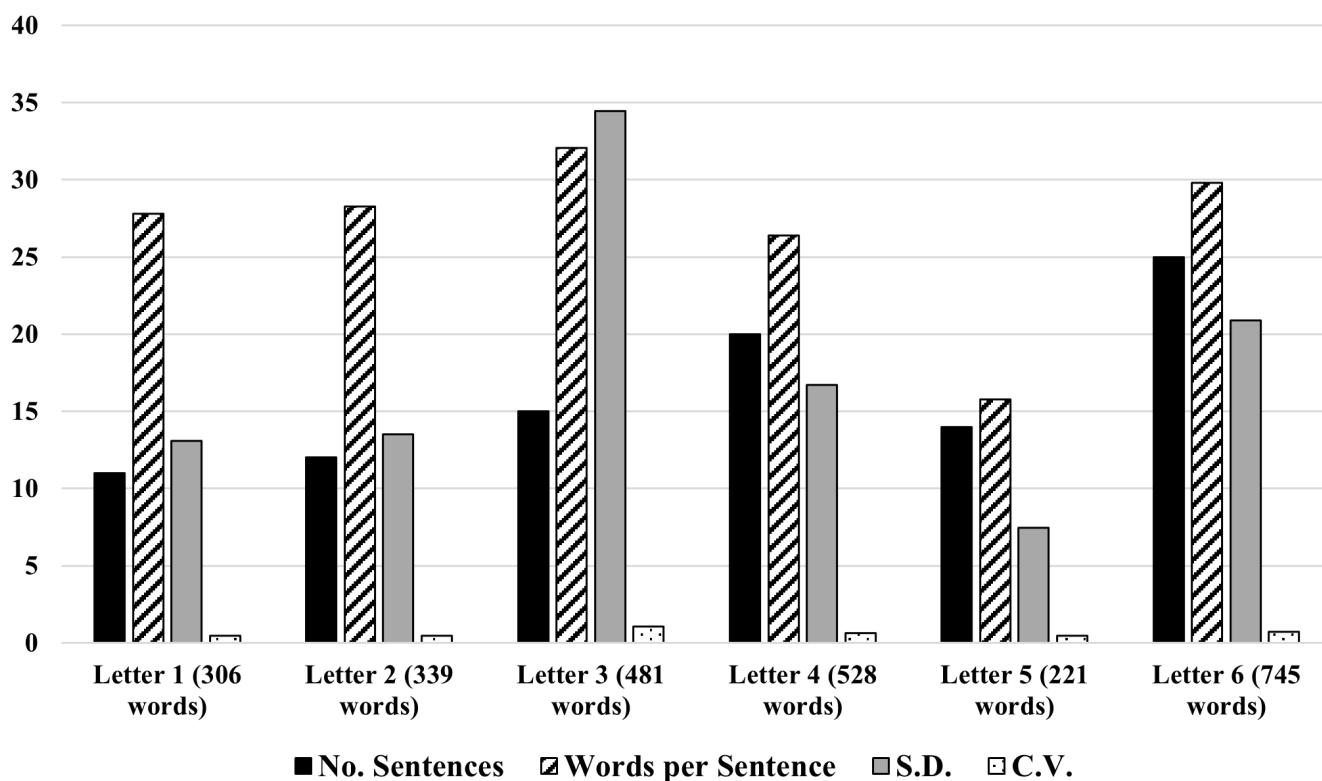
### 3.2.3. A Further Stylometric Approach: Sentence Length

The analysis of sentence length, as practiced in some Forensic Stylometric studies (not by Chaski (2001), though), might in theory also shed some light on authorship, but it weakly does. Table 4 and Figure 5 show the number of sentences, the means of words per sentence, the standard deviation (S.D.) and coefficient of variation (C.V.) in Margery's six letters as measurements of central tendency, which may provide us with variability and differences in distribution (Cantos 2013: 2–10). Letter 3 has the widest sentence length,

as revealed by its means (32.06), standard deviation (34.45) and coefficient of variation (1.07) indexes – which also means that the distribution of variability in this letter is also the highest and least homogeneous. Letter 6 has the second highest number of words per sentence ( $\bar{x}$ =29.80; S.D.=20.89) but also has the highest number of sentences (25). Letter 5 has the shortest sentence length ( $\bar{x}$ =15.78; S.D.=7.44) although a relatively high number of sentences (14) considering the small number of words in the letter (221), which means that its distribution of variability is least and more homogeneous. Given that these patterns do not match those found in spelling or even in grammatical practices, sentence length analysis does not seem to be relevant here for authorship elucidation – probably because, as grammatical features, it is under the strong influence of dictation.

**Table 4: Length sentence in Margery's Letters.**

Letter	Date	No. Words	No. Sentences	Words per Sentence	S.D.	C.V.
<b>Letter 1: #417</b>	1477	306	11	27.81	13.09	0.47
<b>Letter 2: #418</b>	1477	339	12	28.25	13.49	0.47
<b>Letter 3: #419</b>	1481	481	15	32.06	34.45	1.07
<b>Letter 4: #420</b>	1481	528	20	26.4	16.69	0.63
<b>Letter 5: #421</b>	1486	221	14	15.78	7.44	0.47
<b>Letter 6: #422</b>	1489	745	25	29.80	20.89	0.70



**Figure 5: Length sentence in Margery's Letters.**

#### 4. Conclusion

When dealing with authorship in written documents, Historical Sociolinguistics and Forensic Linguistics speak the same language, as they share similar interests. They can thus mutually benefit each other, by sharing their expertise in authorship research and its application to current and historical texts in their social context. The use of the two main approaches to authorship attribution mentioned above, applying Forensic Linguistics techniques to materials from the past can complementarily provide Historical Sociolinguistics with crucial forensic information, and, in turn, put their own theoretical tenets and techniques to the test of time.

In the case of Historical Sociolinguistics, the compilation of large electronic corpora is allowing scholars to carry out complex sociolinguistic analyses. However, the high levels of illiteracy among female population and the subsequent use of scribes make authorship constitute some of the most controversial issues when doing socio-historical research. As Bergs (2005; 2015) suggested, dictation would transmit the author's morphosyntactic or lexical features, but not phonological or graphological ones, which would

be part of the usual sociolinguistic practice of the scribe. For this reason, in periods of a strong gender-based illiteracy, great care has to be exercised in interpreting socio-demographic patterns of variation in epistolary documents. This also has implications for a theory of authorship attribution in Forensic Linguistics. The data suggests that, even when scribes were involved, lexicogrammatical style markers remained stable and most probably reflected the speech of the author, as claimed by Chaski (2001). Crucially, this does not mean that we cannot obtain valuable confirmation for our findings using Forensic Stylistic approaches (data-driven). In spite of their apparently contradictory nature, both trends of research into authorship could be complementary, particularly when applying these to sociohistorical data. In this context, we are not constrained by the severe conditions required to accept forensic evidence in court. Forensic Stylometrics (research-driven and heavily statistical in nature) can go hand in hand with Forensic Stylistics (data-driven, often qualitative) to help us in the elucidation of authorship in historical texts. **N**

JUAN A. CUTILLAS-ESPINOSA & JUAN M. HERNÁNDEZ-CAMPOY

UNIVERSIDAD DE MURCIA

## References

- AMMON, Ulrich, Klaus J. Mattheier & Peter H. Nelde (eds.) 1999. *Historische Soziolinguistik*. Tübingen: Niemeyer.
- ARCHER, Rowena E. 1992. 'How ladies ... who live on their manors ought to manage their households and states': *Women as landholders and administrators in the later Middle Ages. Woman is a Worthy Wight: Women in English Society. C. 1200–1500*, ed. P.J.P. Goldberg. 149–181. Wolfeboro Falls, NH: Alan Sutton Publishing Inc.
- AUER, Anita 2015. Stylistic variation. *Letter Writing and Language Change*, eds. Anita Auer, Daniel Schreier & Richard Watts. 133–155. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139088275.009>
- AUER, Anita, Daniel Schreier & Richard Watts (eds.) 2015. *Letter Writing and Language Change*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139088275>
- AYRES-BENNETT, Wendy 2001. Socio-historical linguistics and the history of French. *Journal of French Language Studies* 11: 159–177. <https://doi.org/10.1017/S0959269501000217>
- BAKER, Paul 2010. *Sociolinguistics and Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- BALDWIN, Thomas Whitfield 1943. *William Shakespeare's Petty School*. Urbana: University of Illinois Press.
- BALDWIN, Thomas Whitfield 1944. *William Shakespeare's Smalle Latine & Lesse Greeke* (2 vols.). Urbana: University of Illinois Press.
- BARRATT, Alexandra 1992/2010. *Women's Writing in Middle English*. London & New York: Longman.
- BAUER, Laurie 2002. Inferring Variation and Change from Public Corpora. *The Handbook of Language Variation and Change*, eds. Jack K. Chambers, Peter J. Trudgill & Natalie Schilling-Estes. 97–114. Oxford: Blackwell. <https://doi.org/10.1111/b.9781405116923.2003.00007.x>
- BELL, Allan 1984. Language style as Audience Design. *Language in Society* 13: 145–204. <https://doi.org/10.1017/S004740450001037X>
- BENSKIN, Michael 1977. Local archives and Middle English dialects. *Journal of the Society of Archivists* 5(8): 500–514. <https://doi.org/10.1080/00379817709514090>
- BENSKIN, Michael 1982. The letters <p> and <y> in later Middle English, and some related matters. *Journal of the Society of Archivists* 7: 13–30. <https://doi.org/10.1080/00379818209514199>
- BERGS, Alexander 2005. *Social Networks and Historical Sociolinguistics: Studies in Morphosyntactic Variation in the Paston Letters (1421–1503)*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110923223>
- BERGS, Alexander 2007a. Spoilt for choice? <The> problem <pe> in <De> Peterborough Chronicle. *The Language of the Peterborough Chronicle*, eds. Alexander Bergs & Janne Skaffari. 45–56. Bern: Peter Lang.
- BERGS, Alexander 2007b. Letters: A new approach to text typology. *Letter Writing*, eds. Terttu Nevalainen & Sanna-Kaisa Tanskanen. 27–46. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/bct.1.04ber>
- BERGS, Alexander 2012. The Uniformitarian Principle and the risk

- of anachronisms in language and social history. *The Handbook of Historical Sociolinguistics*, eds. Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre. 80–98. Malden: Wiley-Blackwell. <https://doi.org/10.1002/9781118257227.ch5>
- BERGS, A. 2015 Linguistic fingerprints of authors and scribes. *Letter Writing and Language Change*, eds. Anita Auer, Daniel Schreier & Richard Watts. 114–132. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139088275.008>
- BIBER, Douglas 1995. *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511519871>
- BIBER, Douglas 2001. Dimensions of variation among eighteenth century registers. *Towards a History of English as a History of Genres*, eds. Hans-Jurgen Diller & Manfred Gurlach. 89–110. Heidelberg: Carl Winter.
- BIBER, Douglas & Edward Finegan 1989. Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text* 9: 93–124. <https://doi.org/10.1515/text.1.1989.9.1.93>
- BIBER, Douglas & Edward Finegan 1997. Diachronic relations among speech-based and written registers in English. *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*, eds. Terttu Nevalainen & Leena Kahlas-Tarkka. 253–275. Helsinki: Societe Neophilologique.
- BLAKE, Norman 1992. Introduction. *The Cambridge History of the English Language*. Vol. 2: 1066–1476, ed. Norman Blake. 1–22. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CHOL9780521264754.002>
- BLAKE, Norman 1996. *A History of the English Language*. London: MacMillan. <https://doi.org/10.1007/978-1-349-24954-1>
- BLUME, Richard 1882. *Die Sprache der Paston Letters. Beitrag zur historischen Grammatik des Englischen*. Programm der Realschule beim Doventhor zu Bremen.
- BRINTON, Laurel J. & Leslie K. Arnovick 2006/2011. *The English Language. A Linguistic History*. Ontario: Oxford University Press.
- CANTOS, Pascual 2012. The use of linguistic corpora for the study of linguistic variation and change: Types and computational applications. *The Handbook of Historical Sociolinguistics*, eds. Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre. 99–122. Malden: Wiley-Blackwell. <https://doi.org/10.1002/9781118257227.ch6>
- CANTOS, Pascual 2013. *Statistical Methods in Language and Linguistic Research*. Sheffield: Equinox Publishing.
- CANTOS, Pascual & Sánchez, Aquilino 2000. Introduction. *Cuadernos de Filología Inglesa* 9(1): 1–3.
- CASTOR, Helen 2004. *Blood and Roses*. The Paston Family in the Fifteenth Century. London: Faber and Faber.
- CHAMBERS, Jack K. & Peter Trudgill 1980. *Dialectology*. Cambridge: Cambridge University Press.
- CHASKI, Carole 2001. Empirical evaluations of language-based author identification techniques. *Forensic Linguistics* 8(1): 1–65. <https://doi.org/10.1558/sll.2001.8.1.1>
- CHASKI, Carole 2005. Who's at the keyboard? Authorship attribution in digital evidence investigations. *International Journal of Digital Evidence* 4(1): 1–13.

- CHASKI, Carole 2007. The keyboard dilemma and authorship identification. *Advances in Digital Forensics III*, eds. Philip Craiger & Sujeet Sheno. 133–146. New York, NY: Springer. [https://doi.org/10.1007/978-0-387-73742-3\\_9](https://doi.org/10.1007/978-0-387-73742-3_9)
- CHEREWATUK, Karen & Ulrike Wiethaus 1993. *Dear Sister: Medieval Women and the Epistolary Genre*. Philadelphia: University of Pennsylvania Press.
- CONDE-SILVESTRE, Juan Camilo 2016. A ‘third-wave’ historical sociolinguistic approach to late Middle English correspondence: Evidence from the Stonor Letters. *Current Trends in Historical Sociolinguistics*, ed. Cinzia Russi. 46–66. Warsaw/Berlin: Open De Gruyter. <https://doi.org/10.1515/9783110488401-006>
- CONDE-SILVESTRE, Juan Camilo & Juan Manuel Hernández-Campoy 2004. A sociolinguistic approach to the diffusion of Chancery written practices in late fifteen century private correspondence. *Neuphilologische Mitteilungen* 105: 133–152.
- CONDE-SILVESTRE, Juan Camilo & Juan Manuel Hernández-Campoy 2012. Introduction. *The Handbook of Historical Sociolinguistics*, eds. Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre. 1–8. Malden: Wiley-Blackwell. <https://doi.org/10.1002/9781118257227.ch>
- CONDE-SILVESTRE, Juan Camilo & Hernández-Campoy, Juan Manuel 2013. Tracing the generational progress of language change in fifteenth century English: the diffusion of <th> in the Paston Letters. *Neuphilologische Mitteilungen* 114(3): 279–299.
- COULTHARD, Malcolm 1994. On the use of corpora in the analysis of forensic texts. *Forensic Linguistics* 1(1): 27–43. <https://doi.org/10.1558/ijsl.v1i1.27>
- COULTHARD, Malcolm 2004. Author identification, idiolect, and linguistic uniqueness. *Applied Linguistics* 25(4): 431–447. <https://doi.org/10.1093/applin/25.4.431>
- COULTHARD, Malcolm & Alison Johnson 2007. *An Introduction to Forensic Linguistics: Language in Evidence*. Oxon: Routledge.
- COULTHARD, Malcolm & Alison Johnson 2010. *The Routledge Handbook of Forensic Linguistics*. London: Routledge. <https://doi.org/10.4324/9780203855607>
- CRAWFORD, Patricia 1985. Women’s Published Writings 1600–1700. *Women in English Society 1500–1700*, ed. Mary Prior. 211–282. London & New York: Methuen.
- CRESSY, David 1980. *Literacy and the Social Order: Reading and Writing in Tudor and Stuart England*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511560484>
- CRESSY, David 1981. Levels of illiteracy in England 1530–1730. *Literacy and Social Development in the West: A Reader*, ed. Harvey J. Graff. 105–124. Cambridge: Cambridge University Press.
- CRESSY, David 1993. Literacy in context: Meaning and measurement in Early Modern England. *Consumption and the World of Goods*, eds. John Brewer & Roy Porter. 305–319. London: Routledge,.
- CRYSTAL, David 2012. *Spell It Out: The Singular Story of English Spelling*. London: Profile Books.
- DAVIS, Gwenn & Beverly A. Joyce 1989. *Personal Writings by Women to 1900: A Bibliography of American and British Writers*. London: Mansell.
- DAVIS, Norman 1954. The language of the Pastons. *Middle English Literature:*

- British Academy Gollancz Lectures*, ed. John Anthony Burrow. 45–70. Oxford: Oxford University Press.
- DAVIS, Norman 1965. That language of the Pastons. *Proceedings of the British Academy* 40: 119–144.
- DAVIS, Norman 1967. Style and stereotype in early English letters. *Leeds Studies in English* 1: 7–17.
- DAVIS, Norman (ed.) 1971. *Paston Letters and Papers of the Fifteenth Century*. Oxford: Clarendon.
- DAVIS, Norman 1983. The language of two brothers in the fifteenth century. *Five Hundred Years of Words and Sounds. A Festschrift for Eric Dobson*, eds. Eric G. Stanley & Douglas Gray. 22–28. Cambridge: D.S. Brewer.
- DAYBELL, James (ed.) 2001. *Early Modern Women's Letter Writing, 1470–1700*. Basingstoke: Palgrave. <https://doi.org/10.1057/9780230598669>
- DOSSENA, Marina & Gabriella Del Lungo Camiciotti (eds.) 2012. *Letter Writing in Late Modern Europe*. Amsterdam & Philadelphia: John Benjamins. <https://doi.org/10.1075/pbns.218>
- DOSSENA, Marina & Susan Fitzmaurice (eds.) 2006. *Business and Official Correspondence: Historical Investigations*. Bern: Peter Lang.
- DOSSENA, Marina & Ingrid Tieken-Boon van Ostade (eds.) 2008. *Studies in Late Modern English correspondence: Methodology and Data*. Bern: Peter Lang.
- DUCKETT, Eleanor 1965. *Women and their Letters in the Early Middle Ages*. Baltimore, Maryland: Barton-Gillet.
- ELSPASS, Stephan 2002. Alter Wein und neue Schlauche? Briefe der Wende zum 20. Jahrhundert und Texte der neuen Medien – ein Vergleich. *Briefkultur im 20. Jahrhundert*, eds. U. Schmitz & E.L. Wyss. 7–31. Duisburg: Gilles & Francke.
- ELSPASS, Stephan 2012. The use of private letters and diaries in sociolinguistic investigation. *The Handbook of Historical Sociolinguistics*, eds. Juan M. Hernández-Campoy & Juan Camilo Conde-Silvestre. 156–169. Malden: Wiley-Blackwell.
- FINKE, Laurie 1999. *Women's Writing in English: Medieval England*. London: Longman.
- GOLDBERG, Peter Jeremy Piers (ed.) 1992. *Woman is a Worthy Wight: Women in English Society. C. 1200–1500*. Wolfeboro Falls, NH: Alan Sutton Publishing Inc.
- GRANT, Tim 2007. Quantifying evidence in forensic authorship analysis. *The International Journal of Speech, Language and the Law* 14(1): 1–25. <https://doi.org/10.1558/ijsll.v14i1.1>
- GRANT, Tim 2010. Txt 4n6: Idiolect free authorship analysis? *The Routledge Handbook of Forensic Linguistics*, eds. Malcolm Coulthard & Alison Johnson. 508–522. Abingdon & New York: Routledge.
- GRANT, Tim & Kevin L. Baker 2001. Identifying reliable, valid markers of authorship: A response to Chaski. *Forensic Linguistics* 8(1): 66–79. <https://doi.org/10.1558/sll.2001.8.1.66>
- HERNÁNDEZ-CAMPOY, Juan Manuel 2008. Overt and covert prestige in Late Middle English: A case study in East Anglia. *Folia Linguistica Historica* 29: 1–26. <https://doi.org/10.1515/FLIH.2008.1>
- HERNÁNDEZ-CAMPOY, Juan Manuel 2013. Ladylikeness and sociolinguistic submission in late medieval English society: Gender-based use of negation in John Paston I and Margaret Paston. *Atlantis* 35 (1): 11–34.
- HERNÁNDEZ-CAMPOY, Juan Manuel 2016a. *Sociolinguistic Styles*. Malden: Wiley-



- Blackwell. <https://doi.org/10.1002/9781118737606>
- HERNÁNDEZ-CAMPOY, Juan Manuel 2016b. Authorship and gender in English historical sociolinguistic research: Samples from the Paston Letters. *Current Trends in Historical Sociolinguistics*, ed. Cinzia Russi. 108–142. Warsaw/Berlin: Open De Gruyter. <https://doi.org/10.1515/9783110488401-009>
- HERNÁNDEZ-CAMPOY, Juan Manuel & Conde-Silvestre, Juan Camilo 1999. The social diffusion of linguistic innovations in 15th century England: Chancery spellings in private correspondence. *Cuadernos de Filología Inglesa* 8: 251–274.
- HERNÁNDEZ-CAMPOY, Juan Manuel & Conde-Silvestre, Juan Camilo (eds.) 2012. *The Handbook of Historical Sociolinguistics*. Malden: Wiley-Blackwell. <https://doi.org/10.1002/9781118257227>
- HERNÁNDEZ-CAMPOY, Juan Manuel & Conde-Silvestre, Juan Camilo 2015. Variability and change in early epistolary English documents: Towards a reconstruction of sociolinguistic change in the history of English. *Letter Writing and Language Change*, eds. Anita Auer, Daniel Schreier & Richard Watts. 14–34. Cambridge: Cambridge University Press.
- HERNÁNDEZ-CAMPOY, Juan Manuel & García-Vidal, Tamara 2018a. Persona management and identity projection in English medieval society: Evidence from John Paston II. *Journal of Historical Sociolinguistics* 4(1): 1–31. <https://doi.org/10.1515/jhsl-2016-0027>
- HERNÁNDEZ-CAMPOY, Juan Manuel & García-Vidal, Tamara. 2018b. Style-shifting and accommodative competence in late Middle English written correspondence: Putting Audience Design to the test of time. *Folia Linguistica Historica* 39(2): 383–420. <https://doi.org/10.1515/flih-2018-0014>
- HERNÁNDEZ-CAMPOY, Juan Manuel & Natalie Schilling-Estes 2012. The application of the quantitative paradigm to historical sociolinguistics: Problems with the Generalizability Principle. *The Handbook of Historical Sociolinguistics*, eds. Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre. 63–79. Malden: Wiley-Blackwell. <https://doi.org/10.1002/9781118257227.ch4>
- HOPE, Jonathan 1994. *The Authorship of Shakespeare's Plays: A Sociolinguistic Study*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511518942>
- HOUSTON, R.A. 1985. *Scottish Literacy and the Scottish Identity. Illiteracy and Society in Scotland and Northern England 1600–1800*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511522598>
- HOWALD, Blake Stephen 2008. Authorship attribution under the rules of evidence: empirical approaches – a layperson's legal system. *The International Journal of Speech, Language and the Law* 15(2): 219–247. <https://doi.org/10.1558/ijssl.v15i2.219>
- JAHR, Ernst H. (ed.) 1999. *Language Change. Advances in Historical Sociolinguistics*. Berlin & New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110807653>
- JENSEN, Vibeke 2012. The consonantal element (th) in some late Middle English Yorkshire texts. *Outposts of Historical Corpus Linguistics: From the Helsinki Corpus to a Proliferation of Resources. Studies in Variation, Contacts and Change in English* vol. 10, eds. Jukka Tyrkks et al. Helsinki:

- Research Unit for Variation, Contact and Change in English.
- KASTOVSKY, Dieter & Arthur Mettinger (eds.) 2000. *The History of English in a Social Context. Essays in Historical Sociolinguistics*. Berlin & New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110810301>
- KROGULL, Andreas, Gijsbert Rutten & Marijke van der Wal 2017. Relativisation in Dutch diaries, private letters and newspapers (1770–1840): A genre-specific national language? *Exploring Future Paths for Historical Sociolinguistics*, eds. Tanja Säily et al. 157–186. Amsterdam & Philadelphia: John Benjamins. <https://doi.org/10.1075/ahs.7.07kro>
- KRUG, Rebecca 2002. *Reading Families: Women's Literate Practice in Late Medieval England*. Ithaca: Cornell University Press.
- LABOV, William 1966. *The Social Stratification of English in New York City*. Washington D.C.: Center for Applied Linguistics.
- LABOV, William 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- LABOV, William 1994. *Principles of Linguistic Change I: Internal Factors*. Oxford: Blackwell.
- LASS, Roger 1992. Phonology and morphology. *The Cambridge History of the English Language. Vol 2: 1066–1476*, ed. Norman Blake. 23–156. Cambridge: Cambridge University Press.
- LASS, Roger 1994. *Old English. A Historical Linguistic Companion*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511621000>
- LASS, Roger (ed.) 2000. *The Cambridge History of the English Language III: 1476–1776*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CHOL9780521264761>
- MCMENAMIN, Gerald 1993. *Forensic Stylistics*. Amsterdam: Elsevier.
- MCMENAMIN, Gerald 2001. Style markers in authorship studies. *The International Journal of Speech, Language and the Law* 8: 93–97. <https://doi.org/10.1558/sll.2001.8.2.93>
- MCMENAMIN, Gerald 2002. *Forensic Linguistics – Advances in Forensic Stylistics*. Boca Raton, FL: CRC Press. <https://doi.org/10.1201/9781420041170.ch9>
- MCMENAMIN, Gerald 2004. Disputed authorship in U.S. law. *The International Journal of Speech, Language and the Law* 11(1): 73–82. <https://doi.org/10.1558/sll.2004.11.1.73>
- MCMENAMIN, Gerald 2010. Theory and practice of Forensic Stylistics. *The Routledge Handbook of Forensic Linguistics*, eds. Malcolm Coulthard & Alison Johnson. 487–507. Abingdon & New York: Routledge.
- MILLWARD, Celia M. 1989. *A Biography of the English Language*. New York: Holt, Rinehart and Winston.
- MILROY, James. 1992. *Linguistic Variation and Change: On the Historical Sociolinguistics of English*. Oxford: Blackwell.
- MOORE, Samuel & Albert H. Marckwardt 1981. *Historical Outlines of English Sounds and Inflections*. Ann Arbor: George Wahr.
- MORTON, Andrew Q. & Sidney Michaelson 1990. *The Q-SUM Plot*. Internal report CSR-3-90, Department of Computer Science, University of Edinburgh.
- MOSTELLER, Frederick & David L. Wallace 1964. *Inference and Disputed Authorship. The Case of the Federalist Papers*. New York: Springer-Verlag.
- MURPHY, Bróna 2010. *Corpus and Sociolinguistics: Investigating Age and Gender in Female Talk*. Amsterdam &

- Philadelphia: John Benjamins. <https://doi.org/10.1075/scl.38>
- NEVALA, Minna & Minna Palander-Collin 2005. Letters and letter writing: Introduction. *European Journal of English Studies* 9(1): 1–7. <https://doi.org/10.1080/13825570500067903>
- NEVALAINEN, Terttu 1999. Making the best use of ‘bad’ data. Evidence for sociolinguistic variation in Early Modern English. *Neuphilologische Mitteilungen* 104: 499–533.
- NEVALAINEN, Terttu & Helena Raumolin-Brunberg (eds.) 1996. *Sociolinguistics and Language History: Studies Based on the Corpus of Early English Correspondence*. Amsterdam: Rodopi.
- NEVALAINEN, Terttu & Helena Raumolin-Brunberg 1998. Reconstructing the social dimension of diachronic language change. *Historical Linguistics 1995. Vol II: Germanic Linguistics*, eds. Richard M. Hogg & Linda van Berger. 189–209. Amsterdam & Philadelphia: John Benjamins. <https://doi.org/10.1075/cilt.162.15nev>
- NEVALAINEN, Terttu & Helena Raumolin-Brunberg 2003/2017. *Historical Sociolinguistics. Language Change in Tudor and Stuart England*. London: Longman Pearson Education. <https://doi.org/10.4324/9781315475172>
- NEVALAINEN, Terttu & Helena Raumolin-Brunberg 2012. Historical sociolinguistics: origins, motivations, and paradigms. *The Handbook of Historical Sociolinguistics*, eds. Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre. 22–40. Malden: Wiley-Blackwell. <https://doi.org/10.1002/9781118257227.ch2>
- NEVALAINEN, Terttu & Sanna-Kaisa Tanskanen (eds.) 2007. *Letter Writing*. Amsterdam & Philadelphia: John Benjamins. <https://doi.org/10.1075/bct.1>
- NINI, Adrea & Tim Grant 2013. Bridging the gap between stylistic and cognitive approaches to authorship analysis using systemic functional linguistics and multidimensional analysis. *The International Journal of Speech, Language and the Law* 20(2): 173–202. <https://doi.org/10.1558/ijssl.v20i2.173>
- NOBELS, Judith & Marijke van der Wal 2009. Tackling the writer-sender Problem: the new developed Leiden identification procedure (LIP). *Historical Sociolinguistics and Sociohistorical Linguistics* ([http://www.let.leidenuniv.nl/hsl\\_shl/Nobels-Wal.html](http://www.let.leidenuniv.nl/hsl_shl/Nobels-Wal.html)). Last access: 09-03-2021.
- O’DAY, Rosemary 1982. *Education and Society 1500–1800. The Social Foundations of Education in Early Modern Britain*. London & New York: Longman.
- O’MARA, Veronica M. 1996. Female scribal ability and scribal activity in late medieval England: the evidence? *Leeds Studies in English* 27: 87–130.
- OLSSON, John 2004/2008. *Forensic Linguistics*. New York: Continuum.
- OLSSON, John & June Luchjenbroers 2014. *Forensic Linguistics*. London: Bloomsbury.
- ORME, Nicholas 1973. *English Schools in the Middle Ages*. London: Methuen.
- ORME, Nicholas 1984. *From Childhood to Chivalry: The Education of the English Kings and Aristocracy, 1066–1530*. London & New York: Methuen.
- ORME, Nicholas 2006. *Medieval Schools: From Roman Britain to Renaissance England*. New Haven, Conn. & London: Yale University Press.
- PALANDER-COLLIN, Minna 1999. Male and female styles in 17<sup>th</sup> century correspondence: I THINK. *Language Variation and Change* 11(2): 123–141. <https://doi.org/10.1017/S0954394599112018>

- PALANDER-COLLIN, Minna, Minna Nevala & Arja Nurmi 2009. *The language of daily life in the history of English. Studying how macro meets micro. The Language of Daily Life in England (1400–1800)*, eds. Minna Palander-Collin, Minna Nevala & Arja Nurmi. 1–23. Amsterdam & Philadelphia: John Benjamins. <https://doi.org/10.1075/pbns.183.02pal>
- PALANDER-COLLIN, Minna 2010. Correspondence. *The Handbook of Historical Pragmatics*, eds. Andreas H. Jucker & Irma Taavitsainen. 677–703. Berlin & New York: Mouton de Gruyter.
- POWER, Eileen 1922/1964. *Medieval English Nunneries c. 1275 to 1535*. Cambridge: Cambridge University Press.
- PYLES, Thomas & John Algeo 1964/1982. *The Origins and Development of the English Language*. New York: Harcourt and Brace Jovanovich.
- RAUMOLIN-BRUNBERG, Helena 1996. Historical sociolinguistics. *Sociolinguistics and Language History: Studies Based on the Corpus of Early English Correspondence*, eds. Terttu Nevalainen & Helena Raumolin-Brunberg. Amsterdam: Rodopi.
- RICHMOND, Colin 1990/2002. *The Paston Family in the Fifteenth Century (vol. 1): The First Phase*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511560309>
- RICHMOND, Colin 1996. *The Paston Family in the Fifteenth Century (vol. 2): Fastolf's Will*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511582165>
- RICO-SULAYES, Antonio 2011. Statistical authorship attribution of Mexican drug trafficking online forum posts. *The International Journal of Speech, Language and the Law* 18(1): 53–74. <https://doi.org/10.1558/ijssl.v18i1.53>
- ROMAINE, Suzanne 1982. *Socio-Historical Linguistics: Its Status and Methodology*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511720130>
- RUSSI, Cinzia (ed.) 2016. *Current Trends in Historical Sociolinguistics*. Warsaw & Berlin: Open De Gruyter. <https://doi.org/10.1515/9783110488401>
- RUTTEN, Gijsbert & Marijke van der Wal 2014. *Letters as Loot: A Sociolinguistic Approach to Seventeenth and Eighteenth-Century Dutch*. Amsterdam: John Benjamins. <https://doi.org/10.1075/ahs.2>
- SÄILY, Tanja, Arja Nurmi, Minna Palander-Collin & Anita Auer (eds.) 2017. *Exploring Future Paths for Historical Sociolinguistics*. Amsterdam & Philadelphia: John Benjamins. <https://doi.org/10.1075/ahs.7>
- SAIRIO, Anni 2009. *Language and Letters of the Bluestocking Network: Sociolinguistic Issues in Eighteenth-Century Epistolary English*. Helsinki: Société Néophilologique.
- SAIRIO, Anni 2017. *Bluestocking Corpus: Letters of Elizabeth Montagu, 1730s–1780s*. Helsinki: University of Helsinki (<http://bluestocking.ling.helsinki.fi/>).
- SEBBA, Mark & Fligelstone, Steven D. 2001. Corpus Linguistics and Sociolinguistics. *Concise Encyclopedia of Sociolinguistics*, eds. Rajend Mesthrie & R.E. Asher. 765–769. Oxford: Elsevier.
- SCHÄFER, Ursula 1996. The late Middle English Paston letters: A grammatical case in point for reconsidering philological methodologies. *Anglistentag 1995 Greifswald*, eds. Jürgen Klein & Dirk Vanderbeke. 313–323. Tübingen: Niemeyer. <https://doi.org/10.1515/9783111714141-028>

- SCHIEGG, Markus 2016. Code-Switching in Lower-Class Writing: Autobiographies by Patients from Southern German Psychiatric Hospitals (1852–1931). *Journal of Historical Sociolinguistics* 2(1): 47–81. <https://doi.org/10.1515/jhsl-2016-0003>
- SCHNEIDER, Edgar W. 2002. Investigating variation and change in written documents. *The Handbook of Language Variation and Change*, eds. Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes. 67–96. Oxford: Blackwell. <https://doi.org/10.1111/b.9781405116923.2003.00006.x>
- SCRAGG, D.G. 1974. *A History of English Spelling*. Manchester: Manchester University Press.
- SEBBA, Mark & Fligelstone, Steven D. 2001. Corpus linguistics and sociolinguistics. *Concise Encyclopedia of Sociolinguistics*, ed. Rajend Mesthrie. 765–769. Oxford: Elsevier.
- SHAHAR, Shulamith 1983/1990. *The Fourth Estate: A History of Women in the Middle Ages*. London & New York: Routledge.
- SIMON, Joan 1966. *Education and Society in Tudor England*. Cambridge: Cambridge University Press.
- STENROOS, Merja 2004. Regional dialects and spelling conventions in late Middle English. Searches for (th) in LALME data. *Methods and Data in English Historical Dialectology*, eds. Marina Dossena & Roger Lass. 257–285. Bern: Peter Lang.
- STENROOS, Merja 2006. A Middle English mess of fricative spellings: reflections on thorn, yogh and their rivals. *To Make his English Sweete upon his Tonge*, eds. Marcin Krygier & Liliana Sikorska. 9–35. Frankfurt am Maim: Peter Lang.
- STONE, Lawrence 1964. The educational revolution in England, 1560–1640. *Past and Present* 28: 41–80. <https://doi.org/10.1093/past/28.1.41>
- STONE, Lawrence 1969. Literacy and Education in England, 1640–1900. *Past and Present* 42: 69–139. <https://doi.org/10.1093/past/42.1.69>
- SVARTVIK, Jan 1968. *The Evans Statements. A Case for Forensic Linguistics*. Göteborg: University of Gothenburg Press.
- TANABE, Harumi 1999. Composite predicates and phrasal verbs in the *Paston Letters*. *Collocational and Idiomatic Aspects of Composite Predicates in the History of English*, eds. Laurel J. Brinton & Minoji Akimoto. 97–132. Amsterdam & Philadelphia: John Benjamins. <https://doi.org/10.1075/slcs.47.36tan>
- TIEKEN-BOON van Ostade, Ingrid 2005. Of social networks and linguistic influence: the language of Robert Lowth and his correspondents. *International Journal of English Studies* 5(1). 135–157.
- TIEKEN-BOON van Ostade, Ingrid 2006. Eighteenth-century English letters: In search of the vernacular. *Linguistica e Filologia* 21: 113–146.
- TIERSMA, Peter & Lawrence M. Solan 2002. The linguist on the witness stand: forensic linguistics in American courts. *Language* 78: 221–239. <https://doi.org/10.1353/lan.2002.0135>
- TURELL, María Teresa (ed.) 2005. *Lingüística Forense, Lengua y Derecho. Conceptos, Métodos y Aplicaciones*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- TURELL, María Teresa 2010. The use of textual, grammatical and sociolinguistic evidence in forensic text comparison. *The International Journal of Speech, Language and the*

*Law* 17(2): 211–250. <https://doi.org/10.1558/ijssl.v17i2.211>

UPWARD, Christopher & George Davidson 2011. *The History of English Spelling*. Malden: Wiley-Blackwell. <https://doi.org/10.1002/9781444342994>

VAN DER WAL, Marijke & Gijsbert Rutten (eds.) 2013. *Touching the Past: Studies in the Historical Sociolinguistics of Ego-documents*. Amsterdam & Philadelphia: John Benjamins. <https://doi.org/10.1075/ahs.1>

VAN HATTUM, Marije (2017). The language of “Ribbonmen”: A CDA approach to identity construction in nineteenth-century Irish English threatening notices. *Journal of Historical Sociolinguistics* 3(2): 241–262. <https://doi.org/10.1515/jhsl-2017-1007>

VINCENT, David 1989. *Literacy and Popular Culture. England 1750–1914*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511560880>

VOESTE, Anja 2018. The self as a source. A peasant farmer’s letters from prison (1848–1852). *Journal of Historical Sociolinguistics* 4(1): 97–118. <https://doi.org/10.1515/jhsl-2017-0011>

WHEALE, Nigel 1999. *Writing and Society: Literacy, Print and Politics in Britain 1590–1660*. London & New York: Routledge.

WOOD, Johanna L. 2007. Text in context. A critical discourse analysis approach to Margaret Paston. *Letter Writing*, eds. Terttu Nevalainen & Sanna-Kaisa Tanskanen. 47–71. Amsterdam & Philadelphia: John Benjamins. <https://doi.org/10.1075/bct.1.05woo>

WRIGHT, David 2013. Stylistic variation within genre conventions in the Enron email corpus: Developing a text-sensitive methodology for authorship research. *The International Journal of Speech, Language and the*

*Law* 20(1): 45–75. <https://doi.org/10.1558/ijssl.v20i1.45>

ZHANG, Shaomin 2016. Authorship attribution and feature testing for short Chinese emails. *The International Journal of Speech, Language and the Law* 23(1): 71–97. <https://doi.org/10.1558/ijssl.v23i1.20300>