

**Title:** The Rise and Fall of a Change from Below in early modern Spanish: the Periphrasis *Deber De + Infinitive* in Texts of Linguistic Immediacy.

**Abstract:**

This paper deals with the patterns of variation and change undergone by a syntactic variable in early modern Spanish grammar, namely the alternation between *deber* ‘have to, must/should’ and *deber de + infinitive* ‘have to, must/should’ as a modal periphrasis. Based on a 1,500,000 word corpus of immediacy text (private letters, memories) the results of this variationist study suggest that, throughout the 16th century, but more especially during its second half, the prepositional periphrasis gradually became more common, above all in stylistic contexts predominated by the spontaneity and proximity of the relationships between the interlocutors. It was also more frequently found in contexts involving members of northern speech communities, particularly males, the young, and the middle-low social strata, the incidence being especially high at the points where some of these groups intersect. All this suggests a change from below in the Golden Age period, which reached considerable dimensions in a relatively short time, but was destined to stagnate and later decline just as quickly in the centuries that followed as a result of some structural features, such as the special ‘visibility’ of the preposition, which could have led to the stigmatization of the periphrasis in a similar way to what happened in other Spanish syntactic phenomena.

**Keywords:** Historical sociolinguistics, Change from below, Modal Periphrasis, Early modern Spanish, Social factors, Linguistic Variation.

**Short running title:** The Rise and Fall of a Change from Below in Classical Spanish

# The Rise and Fall of a Change from Below in Classical Spanish Syntax: the Periphrasis *Deber De + Infinitive* in Texts of Linguistic Immediacy<sup>1</sup>

## 1. INTRODUCTION

Examples (1) and (2), which have been taken from the corpus used as the basis for this study, illustrate one of the earliest syntactic variation phenomena in the history of the Spanish language, i.e. the alternation between the modal periphrases *deber* + infinitive and *deber de* + infinitive ‘have to, must/should’.

- (1) ... todas esas cosas que bienen guiadas por la mano de dios y a el se le *deben dar* las gracias por todo (*Cartas de particulares en Indias*, 16th century)

‘... all those things that come guided by the Hand of God and it is he to whom thanks should be given for all things ‘ (Letters from private individuals in the Indies, 16th cent.)

- (2) ... y bien pudiera ello ser de q me llebaran preso mas en llegando al rio de sebilla no se que me hiziera q la persona q no tiene omra no *debiera de naser*... (*Cartas de particulares en Indias*, 16th century)

‘... and it could well be that they took me prisoner but on reaching the river in Seville, I do not know what made me think that the person without honour should not be born...’ (Letters from private individuals in the Indies, 16th cent.)

A number of recently published studies concerning this phenomenon have concluded that the variant without the preposition, *deber* + infinitive, has clearly been the most commonly used over the centuries (cf. Yllera 1980, Fernández de Castro 1999). Nevertheless, in early modern Spanish the prepositional variant enjoyed considerable usage, with the highest frequencies of uses in history, although these figures unexpectedly dropped abruptly in the centuries that followed (Balasch 2008, 2012, Blas Arroyo & Porcar 2016). However, in the late 20th century and the beginning of this one, the uses of *deber de* + infinitive seem to rise significantly again, as shown by certain dialectological and sociolinguistic studies conducted on a number of Spanish speech communities (a résumé in Blas Arroyo 2011).<sup>2</sup>

---

<sup>1</sup> This study, conducted within the Research Project ‘Variación y cambio lingüístico a través de textos de inmediatez comunicativa: un proyecto de sociolingüística histórica (Ref. P1-1B2013-01) [Linguistic variation and change through texts of linguistic immediacy: an historical sociolinguistic research project], funded by the University Jaume I (2014-2016).

<sup>2</sup> As pointed out by an anonymous reviewer, the *deber de* + infinitive periphrasis is also alive in American Spanish, where it is especially used in official and learned texts (legal, politics, for instance). Both *deber* + infinitive and *deber de* + infinitive are in synchronic variation from early American Spanish times, even though in Spain, the variation declined in the 18th century (Blas Arroyo & Vellón 2014).

In short, we seem to be dealing with one of those processes of variation and change that never quite reach completion, where what we could call dying variants never succumb to their competitors altogether. In these cases, as the centuries go by, such variants may languish for long periods of history without actually giving up their place to rival forms completely. Indeed, in some cases, they even succeed in rising from their ashes to regain a certain degree of vitality.

One particularly interesting process in this diachronic scenario took place in the Golden Age period, where the prepositional variant not only expanded its range of uses but, at the same time, did so primarily in certain linguistic settings.<sup>3</sup> This process involves some specific grammatical and semantic-pragmatic contexts, such as epistemic meanings (Balasch 2008, 2012) as well as some restricted paradigms, such as past simple forms, negative sentences and contextually intensified statements, as we have shown elsewhere (Blas Arroyo & Porcar 2016).

In this study, this assumed change will be examined in greater depth, by analysing the potential influence of different stylistic and extralinguistic factors with the aim of offering a more complete picture of the change. With regard to the first of these two types of factors, our main interest lies in the potential incidence of certain parameters related to the degree of formality of the texts, such as the topic being dealt with in the text or the relationship between the interlocutors. With respect to the second type, provided that the historical information allows us to do so, our aim is to investigate extralinguistic factors such as historical period, geographic origin of the authors and several classic variables in sociolinguistic studies including gender, social status and the generational cohort of the authors.

As the results of this study suggest, throughout the 16th century, but more especially during its second half and most of the 17th century, *deber de* gradually became more common in stylistic contexts predominated by the spontaneity and proximity of the relationships between the interlocutors. From a dialectal and sociological perspective, it was more often used in contexts involving members of northern speech communities, particularly males, the younger generations and the

---

<sup>3</sup> Initially conceived in the Age of Enlightenment to refer to the Spanish literature of the 16th century, the label *Siglos de Oro* (Golden Age) was later expanded to describe the height of the Spanish culture during the Renaissance of the 16th and the Baroque in the following century. For historians and researchers of the Spanish culture at that time, the liminal moments of this Golden Age period were the publication of the influential *Castilian grammar* by Antonio Nebrija in 1492, and the death of the playwright Calderón de la Barca in 1681.

middle-low social strata, and the incidence was even higher at the points where several of these groups overlap. Taken together, this suggests a change from below in early modern Spanish that reached considerable dimensions in a relatively short space of time, but was destined to stagnate and then decline just as quickly as a result of some of its structural and social features.

All these results will be discussed and the main theoretical implications summarized in sections 4 and 5, respectively. Before that, a brief historical overview of the origin and evolution of the Spanish modal constructions with the verb *deber* will be presented in Section 2. This will be followed by an explanation of the objectives as well as the main methodological tools of the research in Section 3.

## **2. NORMS AND USES OF *DEBER* (*DE*) + *INFINITIVE*: BETWEEN THE FREE VARIATION AND ‘CONFUSION’**

As Yllera (1980: 92) recalls, in classical Latin the verb *debere* initially meant ‘to be a debtor of’ in constructions with nominal direct objects, while in the infinitive the verb meant the same as the noun *deber* ‘duty’ today: ‘*el deber me obliga a*’ [my duty obliges me to]. From these original senses, the uses of *debere* expanded to cover the meanings of obligation or necessity, which are the meanings that it has today in periphrases constructed with the infinitive in all Romance languages, where it competes with other auxiliary verbs such as *haber* and, to a lesser extent, *tener*. As in other Romance languages, the deontic periphrasis with *deber* appears without a preposition in the earliest Castilian texts. However, the possible analogy with other deontic periphrases that do include a preposition, *haber a/de*; *tener a/de*, etc., means that, as far back as the 13th century, one can already find the first cases of alternation between the two verbal constructions. Moreover, from very early times both periphrases also shared epistemic meanings reflecting supposition, belief or conjecture, which were already present in Late Latin and which can also be observed in most of the languages derived from it. Thus, as stressed by Yllera (1980: 128), later attempts by scholars to differentiate between the two in terms of their functions – *deber + inf.* for obligation, necessity and convenience ‘have to, must’, versus *deber de + inf.* for epistemic meanings ‘should’ – could be considered artificial (in the same sense, see Alarcos 2004: 264, Demonte 2001: 92-93). These scholars deny the existence of a functional distinction as it was first put forward almost three centuries ago, when the recently

created *Real Academia Española* (1713) fixed the uses of both forms in its *Diccionario de Autoridades* (1732), the first of the lexicographical works published by the standardising institution. Since then, many Spanish grammarians and normative institutions have defended this functional distinction, especially during the 19th and most of the 20th century (Fernández de Castro 1999).

The studies conducted to date portray this phenomenon of syntactic variability as one of the most paradigmatic cases of distancing between the academic prescriptions and actual usage. For some, this is a consequence of a lack of awareness – and ‘confusion’ – between the true meanings of each periphrasis, while for others it simply confirms the difficulty, or the outright impossibility, of determining what criteria have guided this variability throughout the history of Spanish.

In my opinion, however, this debate is based upon ill-defined premises that were partly the result of the inexistence (until very recently) of studies that explored the full complexity of the variable context of syntactic phenomena such as the one being addressed here. This is examined in the present article, using the hermeneutic tools of historical sociolinguistics and variationism.

### **3. METHODOLOGY**

Before discussing the main results of this study, some of the more relevant implications of the methodology followed in this research will be presented in this section. Section 3.1 addresses some achievements made by Historical Sociolinguistics in studying language variation and change in the last few decades, followed by an explanation of the corpus compiled, made up almost entirely of texts of linguistic immediacy, mainly private letters (Section 3.2). Finally, in Section 3.3, I will present the main objectives as well as the factor groups selected to test them with a logistic regression program in which all these factors are considered at the same time.

#### **3.1. Theoretical and methodological issues in a historical sociolinguistic study**

Since the pioneering study of Romaine (1982) three decades ago, Historical sociolinguistics has placed special emphasis on investigating linguistic changes that have been attested in past stages and the relationship that these changes could have with the socio-historical particularities of the communities under study. Compared with more traditional approaches that account solely for internal factors underlying language changes, several authors have shown the possibility of tracking the extralinguistic

motivations that could lie behind these and which are closely related to the social structure of each historical period (Nevalainen & Raumolin-Brunberg 1996, 2003, Bergs 2005, Conde-Silvestre 2007, Hernández-Campoy & Conde-Silvestre 2012, among others).

The possibilities offered by undertaking studies in historical sociolinguistics in real time can help us to observe the rate of diffusion of changes both in the linguistic and in the social systems, thereby revealing the vigour of those changes, as well as the potential incidents that could affect their success or failure on the diachronic axis. At the same time, dealing with data of considerable time-depth makes it possible to adopt an empirical approach to key topics in sociolinguistics, such as whether language changes take place communally or generationally (Labov 2001: 303ff, Conde-Silvestre 2007: 150). However, although the study of other languages and their varieties has gained in importance over the years (Elspass et al. 2007), historical sociolinguistics research has focused particularly on English, and to a much lesser extent on languages such as Spanish. Hence, studies like the present one are a modest attempt to fill this important gap in the literature of diachronic studies on this latter language.

Along with these insufficiencies, the study of variation and change in its socio-historical context has to deal with some methodological problems and shortcomings. For instance, on analysing the potential influence of different social factors, sociolinguists must always take into account the fact that such factors need to be submitted to a hermeneutic reconsideration in the study of earlier periods in history. Indeed, researchers often have to deal with social categories that are not always easy to delimit and, furthermore, they have to be situated within moments in history in which, as is the case here, such categories do not always match those of the present. Therefore, historical sociolinguistics must not incur in anachronisms and take for granted specifications that are common in sociolinguistics today (Bergs 2012: 80ff).

Another prominent problem has to do with the representativeness of the corpus and the statistical validity of data that, in some cases, are quantitatively scarce (Hernández-Campoy & Schilling 2012: 63ff). Not in vain historical linguistics had to cope for decades with the considerable obstacle of processing data from corpora that are often not very systematic and, in most cases, are simply poor and barely representative. This *bad data* problem, already mentioned by Labov (1972) (see also Nevalainen & Raumolin-Brunberg 2003: 26-27; Hernández-Campoy & Schilling 2012: 64-75), has led historical sociolinguistics to base its more recent studies on huge corpora, such as the

*Corpus of Early English Correspondence* (CEEC) or the *Corpus of Nineteenth-Century English* (CONCE), to cite but a few of the ones that can be considered the most representative and which most studies have been based on in the last few years.

In the Hispanic setting, several mega-corpora have recently been compiled and made available by different institutions and scholars, and they are now starting to be used systematically by researchers working in diachronic linguistics. Yet, some of their characteristics make it difficult to exploit them in sociolinguistic studies like the one being conducted here. Thus, for example, a huge database like the one that the Spanish *Real Academia* first made available to researchers several years ago (*Corpus diacrónico del español* (CORDE) – Diachronic Corpus of Spanish) contains a remarkable number of texts, although the vast majority are literary books and texts from other formal discursive traditions (legal, administrative, historical, scientific, etc.). Additionally, there is the problem of the scarce social variation in the texts, since most of them were written by the social elite and, in quite a large number of cases, with a more than obvious literary intent. Indeed, another mega-corpus like the *Corpus del español* (Davies 2002), offer this greater variety, yet it is only possible to search among different registers for the 20th century, but not for earlier centuries, which means that most of the texts are still of a formal nature.

### **3.2. THE CORPUS: TEXTS OF LINGUISTIC IMMEDIACY FOR THE STUDY OF LINGUISTIC VARIATION IN EARLY MODERN SPANISH**

In order to elude these methodological problems as far as possible, for this historical sociolinguistic study a corpus was compiled ad hoc. The most significant peculiarity of this corpus is the fact that it consists entirely of ego-documents that lie close to the pole of communicative immediacy (Oesterreicher 2004: 729ff), rather than literary works or texts from other formal discursive traditions which linguistics has usually been associated with. The majority of the texts included in the corpus correspond to the genre of private letters, written by individuals from different social and dialectal origins. Nevertheless, to ensure greater cohesion of the data, the analysis was limited to correspondence written by Spaniards, either in Spain or in other countries belonging to the ancient Spanish empire.

Working with collections of private letters has become more and more common in recent times for a number of different reasons. One of the most important reasons is the exceptional historical circumstance afforded by Spain's conquest of America, which

gave rise to a huge volume of letters not only of an official and administrative nature but also from emigrants and relatives who, on both sides of the Atlantic, used them to record their personal stories and dramas, often brimming with affection and feeling (on the growing importance of migrant contexts in historical sociolinguistics research, see Kopaczyk & Jucker 2013). Moreover, attention has also been drawn to the fact that the letters present certain structural characteristics that make them especially attractive for studies on historical variation. It is important to note that they contain many autobiographical details, which make it possible to guess the relationships of power and solidarity between senders and addressees, as well as their social status (Okulska 2010).

Likewise, there is also a huge amount of valuable ethnographic information that enables researchers to unravel some of the details of social life at that time (Raumolin-Brunberg 2005, Dossena & Tieken-Boon van Ostade 2008). Another relevant fact to be taken into account is that the letters were not written with the intention of them ever being published, which ensures that the language employed is closer to the vernacular than in use.<sup>4</sup> It is true that there is no shortage of certain automated linguistic elements in those letters, above all in the headings and endings, and they also contain syntactic traits that, initially, might not be expected in this kind of texts (Cano 1996: 380-396). Nevertheless, both the low socio-cultural level of many of the authors and the considerable degree of affectivity emanating from the letters, which is usually a result of the complicity between the interlocutors, make these texts an increasingly more appealing object for use in historical research. As pointed out by García Mouton (1999: 264) ‘literary texts can imitate reality; letters, despite being subjected to all sorts of moulds, are a part of it’ [translation JLBA].

These letters, many of which are of a private nature, are written by persons of different social strata and dealing with topics and relationships that range from a greater familiarity or intimacy at one end (*Epistolaris d'Hipòlita Roís, Vida y fortuna del emigrante navarro, Al recibo de esta..., Cartas de particulares en Indias, Entre Castro del Río y México, El hilo que nos une, Desde la otra orilla, etc.*), to those of more distant nature of official, administrative or business kind, at the other one (*Epistolario*

---

<sup>4</sup> Nevertheless, most of the so-called *cartas de llamada* [call letters] that are found in the Archivo General de Indias (edited by Otte (1988) and Fernandez Alcaide (2009) were written to present them at the Casa de la Contratación [Contracting House], as a proof that the writer or the receiver had a relative in America. Consequently, although these letters also have a notable degree of immediacy, their writers or senders knew their official use. I am grateful to an anonymous reviewer for making this point.



*del conde de Tendilla, Documentos lingüísticos de la Nueva España, En el nombre del hijo, Cartas de mercaderes, Cartas privadas de Hernando Gorjón, etc.*).

Lastly, a number of other texts belonging to discursive traditions that are also associated with communicative immediacy were added to this epistolary corpus. This is the case of the autobiographies and memoirs of certain humble figures (especially soldiers: *Miguel de Castro, Domingo del Toral, Alonso de Borregán, Ruiz de Arce*), diaries (*Moreno de la Torre*) and testimonies obtained by the Inquisition in the proceedings of the Holy Office against those accused of heresy and other religious crimes (*Conversaciones estrechamente vigiladas*).

The complete list of these texts can be seen in the Appendix. Altogether, the corpus gives voice to more than seven hundred individual from different Spanish regions, and who in different moments wrote -or dictated- letters either from Spain or, in most cases, from the newly conquered territories of America. The total number of words amounted is 1,552,104, of which 55% (N = 858,646) are from the 16th century and the remaining 45% (N = 693,458) are from the following century.

### **3.3 OBJECTIVES AND METHODOLOGY**

From the corpus described above and as part of a broader research project for the analysis of variation and change within the paradigm of modal verbal periphrases in the last five centuries (see footnote 1), in this study the intention is to meet the following objectives:

a) investigate the usage distribution of *deber* and *deber de* + infinitive in texts of communicative immediacy from the 16<sup>th</sup> and 17<sup>th</sup> centuries written by Spaniards, assuming that they are closer to the spoken language than other more formal discursive traditions traditionally used in the diachronic study of Spanish;

b) examine carefully the envelope of the variation at that time by analysing certain stylistic (topic and relationship between the interlocutors) and extralinguistic factors groups (geographic origin, gender, social status, generation) in the selection of these variants and, in particular, of the prepositional one (*deber de*), whose use seems to reach its highest point in early modern Spanish;

c) determine the explanatory hierarchy of these factor groups and the direction of variation inside them, as well as its consistency across the grammar system

d) evaluate the existence of potential changes in the uses of *deber de* throughout the Golden Age period and the influence that these factor groups have on them;

e) explain the reasons that could lie behind some anomalies in the rate of diffusion of these changes, from a sharp increase in the use of *deber de* in a short span of time to an abrupt decline in the period that followed.

Bearing all these objectives in mind, and in agreement with the principle of accountability (Labov 1972), all the occurrences of the two variants in the corpus were selected using a concordance program (Wordsmith 4.0). For the analysis, the prepositional periphrasis (hereafter PP), *deber de* + infinitive, was used as the application value. All the tokens were then coded in accordance with the following factor groups:

a) *Stylistic*: 1) letters of a predominantly family-related and private nature; 2) other (non- personal) letters; and 3) memorial texts (for the rationale underlying all these groups, see Section 6.3 below).

b) *Extralinguistic*: 1) Century (16th, 17th); 2) Period (1st half / 2nd half of the 16<sup>th</sup> and 17<sup>th</sup> centuries); 3) Region (northern regions, central/southern regions); 4) Gender (males/females); 5) Age (youngsters/adults cohorts); 6) Social rank (high/medium/low) (see more details in Section 6.4).

Once all the variants had been coded on a spreadsheet, it was exported into the logistic regression program Goldvarb X. For the sake of clarity, in the following subsections I will analyse the general diachronic results (6.1), followed by a detailed examination of the dialectal (6.2), stylistic (6.3) and extralinguistic factors (6.4).

## **4. RESULTS AND ANALYSIS**

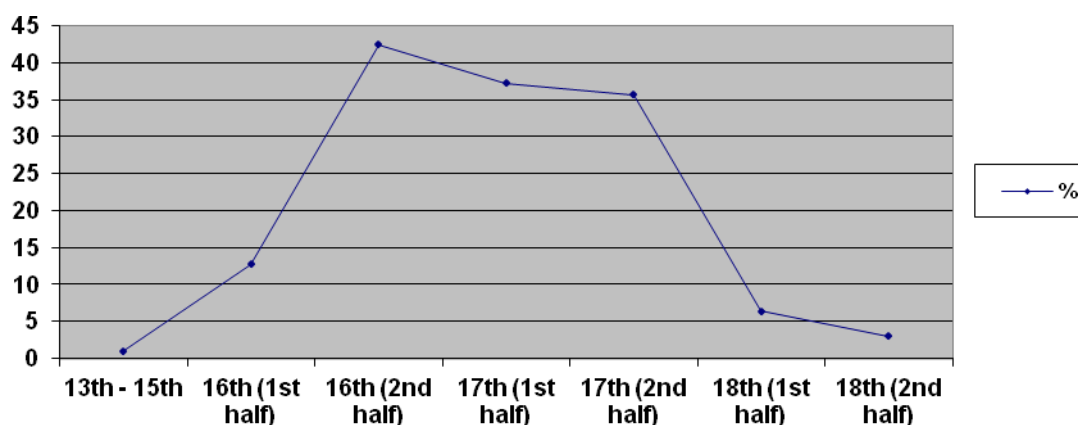
### **4.1. General results and historical evolution of the variable in early modern Spanish**

Altogether the linguistic variable (*deber/deber de*) occurs 547 times, a number that is clearly lower than the one obtained by an alternative periphrasis that predominated at that time (*haber de* + infinitive) (N=2375), but distinctly higher than the combination that is preferred today (*tener que/de* + infinitive) (N=189).

The data initially confirm that the periphrasis with no preposition (NPP) clearly predominates not only in absolute terms (68.4%), but also in the majority of the contexts that were analysed. Thus, the prepositional variant (PP) still enjoyed considerable vitality in the Spanish of the Golden Age period (31.6%), which even seemed to increase in the 17th century, when occurrences of the periphrasis rose to the

far from negligible figure of 37% of the total number, almost 10 points higher than the numbers found in the subcorpus for the previous century (28%). The logistic regression analysis, however, does not select the century as a significant factor because a more thorough analysis of the data shows notable time-related differences within some of those periods.

Indeed, Graph 1 shows how between the 1st and 2nd halves of the 16th century there is a characteristic S-shaped curve that is reminiscent of other *new and vigorous changes* reported in the sociolinguistics literature and which spread forcefully throughout the social matrix (Nevalainen and Raumolin-Brunberg 2003: 130). To judge from the data analysed recently by Balasch (2012: 614), the presence of the PP in medieval texts from the 13th to 15th centuries was just a little over 1%. Nevertheless, the variant seems to awaken in the early 16th century subcorpus (12.7%) in what appears to be an incipient stage of the change (below 15%), and then speeds up substantially (above 40%) towards the end of the century (42.4%). Yet, unlike other similar processes of linguistic change that have ended up becoming consolidated in a relatively short period of time (Ogura & Wang 1996), in our case the PP loses momentum and, after holding its own for some time (first half of the 17th cent.), it began a decline that was quite slow towards the end of the Spanish Baroque period, but picked up speed again in the following century.



**Graph 1: Frequencies of *deber de* in different periods of the history of the Spanish language**  
 (Sources: 13th-15th centuries (Balasch 2012: 614), 18th century (Blas Arroyo & Vellón 2014: 18-19))

A re-analysis of the data into shorter periods, in which the influence of the first part of the 16th century is considered in isolation (.36; 12.7%) against all the others (.61; 39.2%), is now statistically significant (range 25), and suggests the existence of an important shift in favour of the periphrasis *deber de* + infinitive from the second half of the 16th century onwards (see Table 1).

	<b>P.</b>	<b>%</b>	<b>N/Total</b>
<b>Total N (variant):</b> 173/547		31.6	Corrected mean: .281
<b>Century</b>			
16th century		[28]	90/322]
17th century		[37	83/225]
<i>Range</i>	<i>n.s.</i>		
<b>Period</b>			
1st half 16th	.36	12.7	20/157
Rest	.61	39.2	153/390
<i>Range</i>	25		
<b>Region</b>			
Northern	.61	48.1	65/135
Central/southern	.44	27.4	81/296
<i>Range</i>	17		
<b>Tenor</b>			
Family/private letters	.63	42.6	101/237
Other letters	.32	11	20/182
Autobiographical texts	.50	39	50/128
<i>Range</i>	31		
No convergence at iteration: 8			
Log likelihood = -300.241 Significance = 0.009			

**Table 1: Distribution of *deber de* + infinitive by time, dialectal origin and stylistic tenor in a corpus of linguistic immediacy from the 16th and 17th centuries**

However, a more detailed analysis of the data alerts us to the possibility that this eloquent increase is, at least partially, related to several sampling fluctuations that significantly affect some of the most favourable contexts for the PP in the grammar system. In fact, in a previous study (Blas Arroyo & Porcar 2016) we found that the epistemic senses, related to the speaker's judgments of the likelihood of the proposition (Cornillie 2007: 1), as in (3), were one of the most privileged contexts for *deber de* (.79;

61%); in contrast to the deontic ones, dealing with the logic of obligation and necessity (Tagliamonte & Smith 2006: 345), as in (4) (.30; 13%), which have always favoured the alternative periphrasis:

- (3) Estos an sido pliegos que se los trajó la mar porque con quantas diligencias he echo no an parecido y de cansado lo e dexado, No *debe de querer* Dios que sepa lo que passa por allá y así lo llebo en paçiencia (*Desde la otra orilla*, 17th century)

‘These have been documents that were washed up by the sea, because with all the steps I have taken, they have not appeared, and due to fatigue I gave up, God surely does not want me to know what happens there, and thus I bear it patiently’ (From the other shore, 17 th century).

- (4) ... y paréceme ques carta que todos los del Consejo la *debrían de ver* (*En el nombre del hijo*, 16th century)

‘... and it seems to me that it is a letter that everyone in the Council should read (In the name of the father, 16th century)

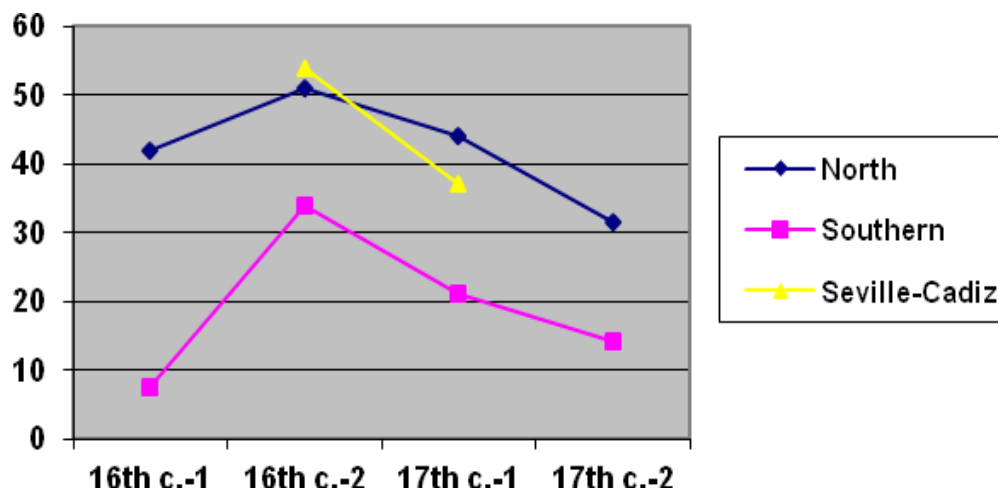
Cross-tabulating these modal meanings and the diachronic axis now shows striking differences in the amount of data available for each context in the different periods. Therefore, if in the first half of the 16th century the non-favourable deontic contexts to the PP are prevalent in the corpus (N=140; 91%), compared to the substantial shortage of the epistemic ones (N=14; 9%), from the second half onwards both modal contexts are now much more balanced, with a significant increase in the epistemic meanings, precisely those more favourable to *deber de* + infinitive (Deontic: N=187; 49% vs. Epistemic: N=195; 51%).

In the light of the historical evolution of Spain throughout that period, might we venture that a country that already finds itself in clear social and economic decline and crisis (Elliott 1998, Fernández Álvarez 2004), as Spain was during the Baroque years, has something to do with the textual hyper-representation of such epistemic contents and therefore with the higher use of the periphrastic variant especially favoured in those contexts? Although our data cannot be conclusive on this point, the hypothesis seems suggestive, an impression that is reinforced by the addition of some complementary and no less striking data.

Indeed, in a similar corpus compiled for the analysis of this variable in the Age of Enlightenment – which, as is well known, was characterised by a faith in the virtues of reason and knowledge – epistemic senses drop drastically again (7.5%), in parallel to the very low uses of the PP in this century (5%) (Blas Arroyo & Vellón 2014: 18-19).

## 4.2 Regional distribution

The dialectal factor is also significant (range 17), and findings suggest that the PP was more characteristic of the regions in the north of the peninsula (north of Madrid) (.61), which are the only ones where the presence of *deber de* practically equalled (48.1%), and that of *deber* (51.9%). Observe also (Graph 2) how in these northern varieties, the PP was already in widespread use in the early 16th century (42%), while its usages increased in the second half of that same century (the true turning point of the linguistic change, as we have seen) to 51%, and lost force again from then onwards from 44% to 31.4%.



Graph 2: Distribution of the frequencies of use of *deber de* in different regions during four historical periods of classical Spanish (%)

As regards the speakers from the central/southern regions of Spain (which are now the autonomous communities of Castilla-La Mancha, Extremadura, Canarias, Murcia and Andalucía), the graph shows a very similar pattern of evolution, although always with considerably lower figures. In the early 16th century, the PP was not at all vigorous (7.6%), it rose sharply again in the second half (33.8%) and finally lost momentum a

short time later (17th c.1st: 21.2%; 17th c.2nd: 14.3%), on what was frankly a decadent path that became generalised in the next century (Blas Arroyo & Vellón 2014).

In light of these data, it seems that the impulse experienced by this modal periphrasis in the classical period could have had something to do with the influence of the northern varieties. In this regard, it is perhaps relevant to note that the highest point in the diffusion of *deber de* coincided with the moving of the Court to Madrid in times of Felipe II (1562), that is to say, during the last few decades of the 16th century, when the S-shaped curve was particularly pronounced (see Graph 1, above). Nonetheless, a hundred years later, the variant receded, perhaps as a consequence of a koineisation process, like the one that has allowed explanations to be put forward for certain well-known phenomena in the history of Spanish, like the devoicing of sibilants (Penny 2000: 42ff;). The end of the *Reconquista* and the ensuing increase in the mobility of individuals and their movement southwards, with the subsequent coexistence of speakers from different social and geographic origins, could explain a process of simplification in which the shorter and less complex variant (*deber* + infinitive) seemed more likely to win than the prepositional one. However, it seems logical that this process would speed up in central/southern regions where the prepositional variant never reached the levels of those in the north of the peninsula.

Now, the letters sent from or to southern Andalusian cities such Seville and Cadiz have deliberately been left out of the catalogue of tokens in this factor. Leaving aside the initial period (for which there are very few data: only six examples evenly distributed between the two variants) and the last one (no data), Graph 2 shows a similar curve to the one described above. Nevertheless, it is quite surprising that the highest point (2nd half of the 16th century) is so high (54%), similar to the one we saw in the northern regions (51%) and considerably higher than the central/southern areas (33.8%). How can this apparently contradictory fact be accounted for on the basis of what has been said up until now?

In my view, these imbalances may be related to the singular role played by these two cities in the trade and immigration that took place with America. Indeed many emigrants, as well as their relatives, had to spend long periods of time in these cities in order to complete all the bureaucratic procedures required of them before they could finally set sail for America (Fernández Alcaide 2009: 53). This means that, in many cases, it becomes impossible to determine where the sender and the addressee were really from. In fact, a comparative analysis between the letters related with immigration,

on the one hand, and other texts (other letters and memorial texts), undoubtedly written by people from these Andalusian areas, offers a very different distributional profile. Thus, the figures obtained by these latter texts are almost half (33%) those related with immigration (52%), which gives us an idea of the need to take those data with precaution.

### **4.3. The stylistic axis of the variability**

In recent years historical sociolinguistics has become increasingly more interested in the analysis of discursive and stylistic factors such as text-types and genres, as well as the growing importance of sociopragmatic notions (i.e. community of practice) for understanding the variability in languages of the past (Włodarczyk 2013: 275-279). To conceptualise some constraints of a basic stylistic nature in this study, two factors are considered: the main tenor of the texts, as well as the degree of relationship between the participants. As a result, we divide the stylistic continuum into three different contexts, in which the parameters of both spontaneity and dialogicality are taken into account:

- 1) Letters of a predominantly family-related and private nature
- 2) Correspondence of an official or commercial nature, which deals mainly with interests that were neither private nor personal
- 3) Memorial texts, in which no addressee is specified and which, therefore, have a lower degree of dialogicality. They may, however, contain important amounts of affectivity, especially in the case of certain ego-documents like autobiographies, memoirs, diaries, etc.

A Goldvarb analysis shows this factor to be highly significant (range 31), and the direction of effects predicted follows these patterns: family letters favour the PP the most (.63) and provide almost one half (42.6%) of all occurrences of the variable. These are followed by the autobiographical texts (.50; 39%), whereas the non-family group of letters clearly disfavours the periphrasis (.32; 11%).

It would seem that the parameters of spontaneity and dialogicality are related with the variation patterns of this syntactic variable: those texts that are located in the most advanced positions on both axes, as is the case of letters dealing with family-related or personal topics, are more likely to make use of the prepositional variant. In contrast, those placed in much lower positions ( official or commercial letters ) favour the PP far less. In fact, some similar cases can be found in the literature on historical sociolinguistics. Thus, Nevalainen and Raumolin-Brunberg (1996: 200) found that the



diffusion of the personal pronoun *its* as a possessive throughout the history of English had an especially favourable initial locus in texts that were closer to the spontaneous and colloquial language (private letters, comedies, diaries or travel books, etc.). In more official texts, however, the occurrences of that variant were initially much less frequent.

How did this spreading of *deber de* throughout the linguistic system take place? To answer this question, the stylistic factor was cross-tabulated with the most significant ones from the linguistic matrix (i.e. epistemic-deontic modal opposition, sentence polarity and contextual modalisation) which have been analysed elsewhere (Blas Arroyo & Porcar 2016).

	P.	%	N
<b>Modal meaning</b>			
Deontic	.30	13.5	44/327
Epistemic	.79	61.2	128/209
<i>Range</i>	49		
<b>Polarity</b>			
Affirmatives	.46	28.9	130/450
Negatives	.66	44.6	41/92
<i>Range</i>	20		
<b>Contextual modalisation</b>			
Intensified	.58	35.8	67/187
No modalised	.44	29.4	106/360
<i>Range</i>	14		

**Table 2: Contributions of the main linguistic factor groups to the selection of *deber de* +infinitive in an immediacy corpus of the 16<sup>th</sup> and 17<sup>th</sup> centuries.**

As can be seen in Table 2, the results of this analysis show that the selection of the PP is significantly associated with the epistemic-deontic modal opposition (range 49), sentence polarity (range 20), and contextual modalisation (range 14). And inside these factor groups, *deber de* is especially favoured by epistemic meanings (.79), sentences of negative polarity (.66), as well as formally intensified contexts (.58).

Now, after cross-tabulating those factor groups with the different types of texts we discover that:

a) The association between the epistemic modality and the prepositional variant is far more pronounced in family letters (69%) and autobiographical texts (74%) than in letters dealing with less personal matters (19%).

b) As for the incidence of negative polarity, this factor is also clearly more relevant in intimate letters (66%) and autobiographical texts (50%) than in letters that do not deal with family-related topics (17%).

c) Finally, a similar pattern is observed in the intersection between the main tenor of the texts and what I have called *contextual modalisation*, which refers to those extracts in the corpus in which the linguistic variable is framed by linguistic resources that represent a formal increase intended to intensify its illocutionary or perlocutionary force. We again find that the modalisation factor is activated mainly in the letters containing more personal and family-related information (57%), closely followed by the autobiographical texts (48%), but is practically neutralised in the non-private missives (15%).

In sum, the results of these cross-tabulations suggest a close correlation between the two kinds of factor groups. Thus, the more strongly marked factors in the broad sphere of modalisation (epistemic meanings, negative polarity and contextually modalised statements) take on considerably greater significance in the selection of PP in texts characterised by a higher degree of emotional involvement on the part of the author. And yet, this significance is attenuated or even reverts in the opposite (non-spontaneous) contexts. Although there are certain differences between some of the cross-tabulations, all of them show that the more neutral, less personal texts, such as public or commercial letters, make less use of the PP. In contrast, in the more intimate and personal letters, as well as to a lesser extent in the autobiographical texts, the joint influence with linguistic modalisation increases sharply.

How do all these patterns of variability behave in the social system? To answer this question, in the following section I analyse the potential influence of different social factors, although this time taking into account *only* the texts in which there is proof that they were either written by the sender or, although handwritten by scribes (a common practice at that time), were dictated directly by the sender or author.<sup>5</sup>

Accordingly, texts that do not offer such proof, as well as copies made at a later date by

---

<sup>5</sup> These data are extracted from the information provided by the corresponding editors of the texts that the corpus includes, which are listed in the Appendix. Nevertheless, when no information of this kind is given, the tokens of the linguistic variable are excluded from this analysis.

individuals who did not participate in the original communicative act, were not coded (on the value of first-hand attestations in the analysis of past stages of languages, see Pahta and Jucker 2011). The results of this new analysis are shown in Table 3.

	<b>P.</b>	<b>%</b>	<b>N/Total</b>
<b>Total N (variant):</b> 173/304		29%	Corrected mean: .273
<b>Century</b>			
16 <sup>th</sup> century		[13.8	20/122]
17 <sup>th</sup> century		[36.3	66/182]
<b>Range</b>	<b>Period</b>		
1 <sup>st</sup> half 16 <sup>th</sup>	.25	5.8	5/86
Rest	.63	37.2	81/218
<i>Range</i>	47		
<b>Region</b>			
Northern	.60	44.1	49/111
Central/southern	.41	19	24/126
<i>Range</i>	19		
<b>Tenor</b>			
Family letters	.64	36.5	27/74
Non-family letters	.17	6.7	7/105
Autobiographical texts	.73	42.7	50/117
<i>Range</i>	56		
<b>Sex</b>			
Male		[29.8	82/275]
Female		[13.8	4/29]
<i>Range</i>	<i>n.s.</i>		
<b>Social rank</b>			
High		[20.1	30/149]
Intermediate		[64.2	43/67]
Common people		[54.1	40/74]
<i>Range</i>	<i>n.s.</i>		
<b>Age</b>			
Younger cohorts		[45.6	36/79]
Adults cohorts		[21.9	37/169]
<i>Range</i>	<i>n.s.</i>		

No convergence at iteration: 7

Log likelihood= -153.042

Significance: 0.001

**Table 3: Contribution of some extralinguistic factors to the selection of the periphrasis *deber de* + infinitive in a corpus of linguistic immediacy texts from the 16th and 17th centuries (NB. Only autographed texts or texts dictated by the author)**

As expected, the number of tokens of the variable is significantly reduced in this second analysis, with a total of 304 occurrences. Nevertheless, it is interesting to note that most of the variability patterns discussed in previous sections are preserved, as shown by comparing Tables 1 and 3. Again, the shift in favour of the PP is seen to speed up from the second half of the 16th century onwards (.63; 37.2%), from the far more testimonial figures recorded in the first half (.25; 5.8%). In contrast, the dialectal profile presents the same distribution plot and a very similar rank (range 19) as the one commented above, with an increase in the use of *deber de* in the northern regions (.60; 44.1%), which is twice the number obtained for the central/southern areas (.41; 19%). Finally, the stylistic axis also conserves most of the conclusions discussed earlier. Hence, the tenor of the texts appears once again as a highly significant factor (range 56), with letters related to non-family matters as one of the contexts that least favours the PP (.17; 6.7%). This is just the opposite of what occurs in the case of letters dealing with family and domestic matters (.64; 36.5%).

These data are interesting because they force us to think about the influence that scribes and other intermediaries involved in setting the text down on paper may have exerted on linguistic varieties that were not their own. If the mark left by these scribes had to play a decisive role in the field of phonic or graphic variation, perhaps it was not so much so in other spheres such as syntax or lexicon (in the same sense, Bergs 2005: 79-80).

#### **4.4. Social factors**

Research on gender in the history of a language can become a nightmare from a methodological point of view (Bergs 2012: 89, Kielkiewicz-Janowiak 2012: 307ff). Prevailing cultural differences meant that, in the best of cases, women knew how to read but could not necessarily write (and in the most popular layers of society more often than not they were unable to do either). This shortage of samples written by women therefore leads to considerable problems in terms of representativeness in historical

sociolinguistic research. In our case this is also true, since occurrences of texts written by female authors account for barely 10% of the data.

Within this small group of female writers, the PP is chosen on very few occasions (N=4; 13.8%), in contrast to what we observe among men (N=82; 29.8%). Apparently, these data do not agree with those of other studies carried out in historical sociolinguistics which seem to display women as leaders in language change regardless of the social embedding of the process (Nevalainen and Raumolin-Brunberg 2003: 130).

Unfortunately, in our case we have little information about female speech, and the few data that are available are spread very unequally over the different periods considered in the study. However, it could be relevant to note that, in the first half of the 17th century, which is when most of the female occurrences are concentrated (23 out of 29 occurrences) and when the spread of *deber de* seems to be consolidated in the community as a whole (see 6.1), women used the PP very sparingly (only three cases; 13%), whereas the figure for men reached a far from negligible 42%. What could be the reasons explaining this lag in the diffusion among women of a variant that underwent an important change in early modern Spanish?

In my view, this fact could largely be explained by the heavy dependence on the gender factor as regards the *social rank* of the population. Indeed, *all* the female occurrences of the variable were the work of people belonging to the uppermost layers of a strongly stratified society like that of Spain in the 16<sup>th</sup> and 17<sup>th</sup> centuries. Hence, the gender variable can only be analysed here in interaction with other important social factors, like *social rank*, in order to determine how variation and change might be conditioned by parameters such as legal rights, access to education or social mobility, which were very different for men and women in Spain at that time (Elliott 1998).

In the social structure which characterised Spain at that time, three main ranks can be distinguished. At the top, we find representatives of the high aristocracy, who held the main titles of the time (dukes, counts, marquis, and so forth) and enjoyed an enormous amount of political, economic and social power due to their close relationships with the monarchy. Likewise, this group also included the elite of the ecclesiastical hierarchy (bishops, archbishops, cardinals, high officials of the Inquisition, superiors in monasteries and convents, etc.), who nearly always came from the noblest of families in the Golden Age period (Elliott 1998, Fernández Álvarez 2004).

A second rank – which cannot really be compared with today's middle classes – is that made up of an assorted list of members of the clergy and lower nobility, whose

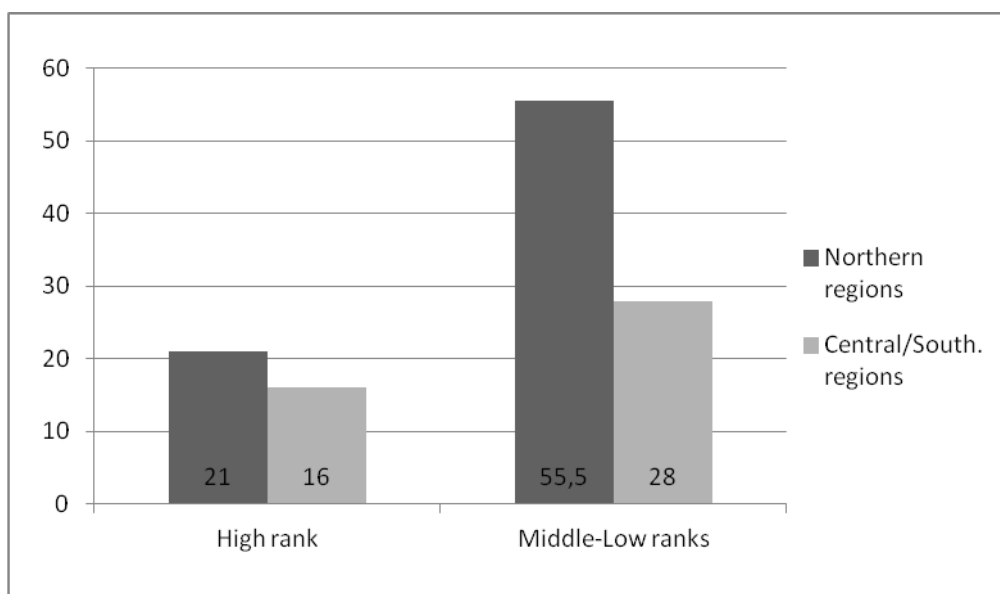
fortune would gradually diminish as time went by. This same group also included representatives of a growing urban bourgeoisie (university graduates, doctors, scribes, merchants, teachers, etc.), who were always at a disadvantage with respect to the aristocracy and the clergy because they were obliged to pay taxes while the latter were exempt from tax pay, which had disastrous effects on the Spanish economy. Lastly, the lowest rank included the common people, who worked in manual occupations, like farmworkers, bakers, carpenters, wood turners, coopers, builders, smiths, and so forth, many of whom also sought their fortune in America and wrote to their relatives with their news or asking their family to join them (Fernández Alcaide 2009: 43-54).

In addition to the different rights and obligations of each social rank, which were often passed on from generation to generation – although there are also a number of significant examples of social ascent (the case of conquerors) and descent (the lower nobility) – there were also differences among the members of those groups in terms of the access they had to education. Hence, for example, the written skills of the language were reserved for the first rank, and were only available to the other two to a far lesser extent. Nevertheless, many female nobles were only able to learn the skill of reading (generally religious books) and writing was an ability that was far less widespread. Indeed, when they were able to write, in the best of cases they could only do so in a faltering, rudimentary manner (Castillo 2006). Logically, there are instances that do not match this description, such as the ones included in this corpus, although they were the exception rather than the rule. Some representatives of the titles and occupations within the intermediate ranks (monks, friars, doctors, traders, noblemen, etc.) were also able to read and write. Such skills were, however, far less universal than among the social elite, which meant that they quite often needed the help of intermediaries to be able to write down what they thought and felt. And for more obvious reasons, this was also the norm among common people, barring the few exceptions that did exist, where some made virtue of necessity and learnt the basics of writing so as to be able to convey news and affection to people thousands of miles away. As a result, the number of examples from these common people remains a considerably long way behind those of the other two.

Although the social rank factor group is not selected as significant by Goldvarb, the differences in frequency between the different social strata speak for themselves, especially those between the low (54.1%) and middle (64.2%) layers of society, on the one hand, and the upper ones (20.1%), on the other. In view of this distribution, it seems that common people in the society of the Spanish Golden Age played a significant role

in the spread of the prepositional variant, in contrast to the elites, who were far more reluctant to use it.

All this suggests a change from below, similar to the one that could have taken place with the devoicing of the Spanish sibilants in the 15th century, which incidentally was also initially disseminated in the northern regions (Penny 2000: 42, 119, Conde-Silvestre 2007: 300-302). In fact, cross-tabulating the social stratification and the dialectal origin of the authors (see Graph 3) shows that it is in the highest ranks where the dialectal differences are less relevant, with values that are quite similar, although slightly higher in the north (21%) than in the centre/south (16%). In contrast, it can be seen how these differences become much greater in the next strata, to the point where the occurrences of *deber de* in the north (55.5%) are twice those of the central/southern territories (28%).



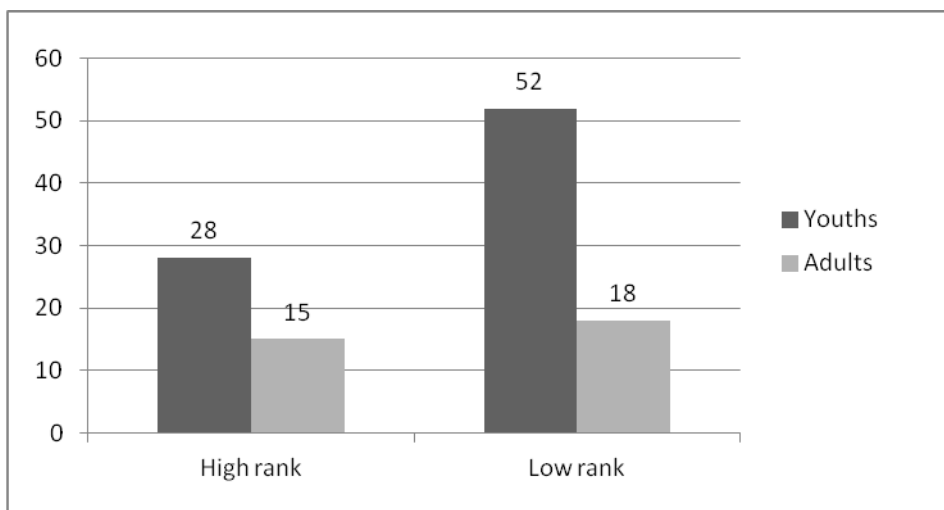
**Graph 3: Distribution of the variant *deber de* + infinitive after the cross tabulation of *region* and *social rank* (%)**

Bearing these numbers in mind, we are now in a better position to compare the behaviour of men and women, even though we have to limit this analysis to the highest rank of the population, since we do not have any female occurrences in the other two. Despite the important differences between the samples of the two groups (Men: 121 occurrences; Women: 29 occurrences), which makes it more difficult to draw statistical inferences, the truth is that the percentages obtained for men and women are now very similar (18% and 14% respectively).

To sum up, these results and the absence of women in the two lower social rank categories lead us to conclude that the low rate of use of the PP variant by women that we highlighted at the beginning of this section may not be a gender effect. Indeed, there is no guarantee that, in a larger corpus, women in lower classes would get higher rates, more like the figures for men. This is undoubtedly an interesting line of research for future investigations in the socio-historical study of Spanish. In order to configure the *generational* factor and given the great difficulties involved in determining exactly the age of many of the speakers were when they wrote their letters and memoirs, we carried out an emic rather than chronological interpretation of age. On this occasion we have therefore resorted to a twofold classification, in which speakers were classified into two groups, depending on what life-events they were experiencing when their texts were written. To do so, we have made use of both what was said in those texts and – especially in the case of letters – indirect information related to the act of communication. Thus, for example, an immigrant who admits to having been in America for only a short time and who is writing to his or her parents, who stayed behind in Spain, is likely to have still been young. In contrast, the father who writes to his son in America is included, for the purposes of our study, as a member of the adult cohorts of society. Occurrences of individuals that could not reasonably be assigned with any direct or indirect data were left uncoded. In the end, there are only 248 tokens available, of which 32% (79) correspond to individuals within the first stage of life versus 68% (169) that could be linked to the adult cohorts.

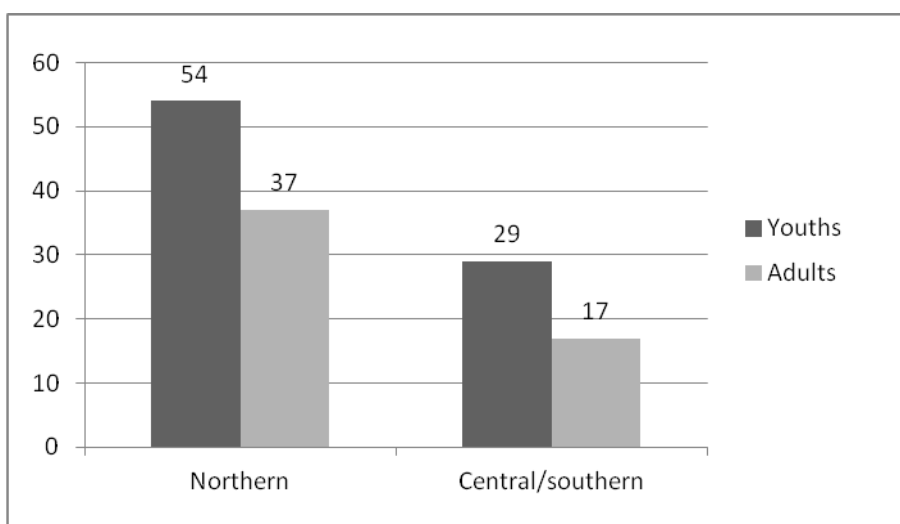
Table 3 shows how the younger authors used the prepositional variant twice as often (45.6%) as the older ones (21.9%). And, interestingly, these differences are also observed both in the upper ranks of society (28% vs. 15%) and in the lower ones, although it is in these latter where the distance between the two subgroups is more pronounced (52% vs. 18%) (see Graph 4).





**Graph 4: Distribution of the variant *deber de* + infinitive after the cross-tabulation of *age* and *social stratification* (%)**

Another promising detail is the intersection between this generational factor and the dialectal axis. In this regard, it is interesting to note that the vitality of *deber de* is particularly noticeable among northern youngsters (see Graph 5), the only subgroup whose realisations of the prepositional variant (54%) outnumber those of the non-prepositional one, and at some distance from the more adult generations (37%). Also in the southern regions, the younger members of the population use the PP (29%) practically twice as much as their elders (17%), which is in line with the fact that the younger generations traditionally adopt foreign novelties more quickly and intensely (Tagliamonte 2012: 43-55).



**Graph 5: Distribution of the variant *deber de* + infinitive after the cross-tabulation of *age* and *region* (%)**

## 5. CONCLUSIONS

From the data analysed so far, it can be deduced that the process of change favouring the spread of *deber de* + infinitive in early modern Spanish was led mainly by the following groups: the members of the northern speech communities, the common people, men and the younger cohorts of the population. Moreover, although the occurrences of the variant at the point where some of these factors intersect are sometimes insufficient due to the sample size, cross-tabulating some of these factors does yield certain regularities that are worthy of analysis. For example, it can be seen how, except at the crossing point with the female gender, where the differences are practically neutralised, in all the other cases the results of those crossings are much higher than those of their alternatives. Thus, it has been observed how, in certain subgroups (i.e. youths-low ranks, low rank-northern), the use of the prepositional variant even exceeds that of the non-prepositional one, which was clearly the predominant periphrasis in society at that time.

Diachronically, the change advanced timidly from the beginning of the 16th century, but it is above all in the second half of that century when it reached its high point, especially in the northern regions of Spain. From that moment on, the change levelled out in the first half of the 17th century, although as the century advanced *deber de* began a slow process of decline that was to speed up drastically in the next century. We are therefore dealing with a process of linguistic change that advanced very swiftly in a relatively short period of time, due to the efforts of certain social groups like those mentioned above. However, instead of consolidating positions, it actually went on to lose them in a process that began slowly but then reached a high speed in the centuries that followed.

What possible explanation can there be underlying this pattern of variation and change, in which we observe a sharp increase in the use of the *deber de* + infinitive variant during the Spanish Golden Age, followed by an abrupt decline in the period immediately afterwards? There is no single answer to this question, I believe, but, as the data from this research have shown, it seems that there is a conjunction of linguistic and extralinguistic reasons behind it. As far as the structural motives are concerned, we have seen elsewhere (Blas Arroyo & Porcar 2016) how the ups and downs of this variant run parallel to the degree to which the contexts that are more favourable to the periphrases are present in the corpus, i.e. the meanings associated with epistemic modality. Hence,

from a timid emergence of these in the first part of the 16<sup>th</sup> century, when the presence of deontic contexts – always favourable to the non-prepositional variant – is overwhelming, such settings rise sharply in the following period, in a scenario that clearly favours *deber de*.

Conversely, in a similar corpus compiled for the analysis of this variable in the 18<sup>th</sup> century, those contexts fall back sharply, exactly like the use of this periphrasis, in a process that seems to have lasted until recently, when some recent data point to a certain revival, at least in peninsular Spanish (cf. Gomez Manzano 1992, Gomez Molina 2009, Blas Arroyo 2011).

Now, are those unexpected changes the consequence of mere chance in sampling, or could they be connected with some key socio-historical moments in the evolution of Spanish society in the Golden Age period? Obviously, the present data set cannot be conclusive in this respect, but maybe it is no coincidence that the episodes in which the meanings of uncertainty, speculation and conjecture predominate in the corpus are those in which –long forgotten the imperial glories of its recent past – Spain is a country in clear social and economic decline. In contrast, the new winds of change brought in by the Age of Enlightenment, with a renewed confidence in the virtues of reason and knowledge as the best way to rescue the nation from the state of depression in which it had been left at the end of the Baroque, coincide in time with a sharp decline in both these meanings and – perhaps not by chance? – the corresponding use of the prepositional periphrasis.

Moreover, these salient variations in the level of use of *deber de* in the Golden Age period could also be related to the dialectal configuration of Spain at that time and the movements of people that characterised the period following the end of the *Reconquista* in the late 15<sup>th</sup> century. In this respect, we have seen how the highest point in the diffusion of this periphrasis, which seems to have always been more popular in the northern regions, coincides with the moving of the Court to Madrid in times of Felipe II (1562). Since then, the variant seems to have spread southwards, although it never enjoyed the success it reached in the north. Furthermore, the increase in mobility of individuals and their movement southwards, with the subsequent coexistence of speakers from different social and geographic origins, could perhaps explain a process of simplification and koineisation in which the less marked variant (*deber*) seemed more likely to win than the prepositional one. However, it seems logical that this

process would speed up in central/southern regions, where the prepositional variant never reached the levels of those in the north of the peninsula.

In contrast, the analysis of the stylistic axis has enabled us to complete the profile of this change in classical Spanish, by observing, for example, how the uses of the prepositional periphrasis are produced at a greater rate in texts of a more family-related and personal nature. At the same time, we have seen how this fact covaries with certain modalising loci in the grammatical system, such as with epistemic meanings, as well as contexts of negative polarity and those characterised by an emphatic nature. At this point, we could ask ourselves: what is the reason for this combined effect? Why does the preposition appear significantly more frequently in the most modalised contexts of the texts where family-related topics and close relations between the interlocutors predominate?

The answer to this question, I believe, lies in the fact that both revolve around the actual concept of modalisation in a broad sense, that is to say, around the attitudes and particularly the affectivity the speaker instils in his or her messages. The speakers might have interpreted the preposition as an expressive resource that allowed them to emphasise their ideas, intentions and desires more effectively. And what better place to do so than in letters where priority is given to affectivity and feelings, in contrast to those in which both the content and tenor are necessarily colder and more distant.

This iconicity of the preposition is distributed unequally across the social and the dialectal axis. It starts to spread with force in the northern regions of the peninsula, from where it moves out to the rest. Yet, unlike other linguistic traits with a similar outcome (the case of the devoicing of the sibilants in the mediaeval period), the social profile of this change from below seems to hold within itself the very seed of its own failure. Therefore, perhaps one of the reasons why the prepositional variant had finally lost prominence in favour of the non-prepositional periphrasis by the end of the Spanish Baroque period was the special ‘visibility’ of this resource (similar, for example, to that observed in other stigmatised phenomena in Spanish, like *dequeísmo*), which is used in more family-related and colloquial contexts, but above all, by part of the social sectors with little, if any, social influence.

But at the same time, *deber de* has the ‘unfortunate knack’ of activating itself to a greater degree in more minority linguistic contexts. Thus, negative polarity or contextual intensification resources appear far less frequently in discourse than their affirmative and non-intensified opposites. And the same happens, as we have seen

earlier, with the epistemic meanings, which are much less frequent in the corpus than the deontic ones, especially in the letters written by immigrants with the aim of urging their relatives to complete the procedures required to be able to unite the family again overseas. This would be another important reason explaining why a process of linguistic change that achieved considerable vigour in classical Spanish was suddenly cut short.

#### **Appendix: list of texts included in corpus**

Ahumada Batlle, Eulàlia, ed. 2003. *Epistolaris d'Hipòlita Roís de Liori i d'Estefania de Requesens*. València: Universitat de València.

Aramburu Zudaire, José Miguel. 1999. *Vida y fortuna del emigrante navarro a Indias, siglos XVI y XVII*. Pamplona: Gobierno de Navarra, Departamento de Educación y Cultura

Arbelo García, Adolfo I. 2010. *Al recibo de esta... Relaciones epistolares canario-americanas del siglo XVII*. Idea: Santa Cruz de Tenerife

Archivo General de Andalucía. 2010. *Un Epistolario de Bernardo José Aldrete (1612-1623)*. Sevilla: Consejería de Cultura.

Baranda, Consolación. 1991. *Cartas entre Felipe IV y Sor María Jesús de Ágreda*, Madrid: Castalia.

Borregan, Alfonso. 2012. *La Conquista del Perú* (ed. by Eva Stoll & María de las Nieves Vázquez Núñez). Madrid: Iberoamericana.

Bouza, Fernando. 2001. Cartas de nuevas y de placer. *Corre manuscrito: una historia cultural del Siglo de Oro* ed. by Fernando Bouza, 170-178; 199-215. Madrid: Marcial Pons.

Carreira, Antonio. 1999. *Luis de Góngora: epistolario completo* (selection of letters). Lausana: Sociedad Suiza de Estudios Hispánicos.

Castillo, Antonio. Dos cartas de Isabel de Baena a la duquesa del Infantado. *Archivo-Ibero-Americano* 1.322-325.

Company, Concepción. 1994. *Documentos lingüísticos de la Nueva España (XVI al XVIII)*. México: El Colegio de México (selection of letters written by Spaniards in the 16 and 17th centuries).

Cortijo, Adelaida & Antonio Cortijo. 2002. Entre Luisa de Carvajal y el conde de Gondomar. Nuevos textos sobre la persecución anticatólica en Inglaterra (1612-1614). *Voz y Letra* 13:2.17-59.

- Cortijo, Adelaida & Antonio Cortijo. 2003. *Cartas desde México y Guatemala (1540-1635. El proceso de Díaz de la Reguera* (Bancroft Library Ms. 92/83z). Cáceres: Universidad de Extremadura.
- Del Toral y Valdés, Domingo. 1905. Relación de la vida del capitán Domingo de Toral y Valdés, escrita por el mismo capitán ed. by Manuel Serrano y Sanz, *Autobiografías y memorias*, 485-506. Madrid: Bailly-Ballière e Hijos.
- Eberenz, Rolf. & Mariela de la Torre. 2002. *Conversaciones estrechamente vigiladas: interacción coloquial y español oral en las actas inquisitoriales de los siglos XV al XVII*. Lausana: Centro de Estudios Hispánicos.
- Fernández Alcalde, Marta. 2010. *Cartas de particulares en Indias del siglo XVI*. Sevilla: Univesidad de Sevilla.
- Fontanella de Weinberg, María. 1993. *Documentos para la historia lingüística de Hispanoamérica*. Madrid: BRAE (selection of letters written by Spaniards in the 16 and 17th centuries).
- Frago, José. A. 1999. Notas sobre el español antillano del siglo XVI en cartas de emigrados andaluces. *Estudios de lingüística hispánica. Homenaje a María Vaquero* ed. by Amparo Morales, **Eduardo Forastieri Braschi, Julia Cardona, Humberto López Morales**, 350-362 San Juan: Editorial de la Universidad de Puerto Rico.
- Gómez Seibane, Sara, Carmen Isasi Martínez & Enriqueta Sesnero Cutanda. 2007. *Bilbao en sus documentos (1544-1694)*. Deusto: Universidad de Deusto.
- González, Jesús. 1999. *Epistolario y poesías / Luisa de Carvajal y Mendoza*. Madrid: Atlas.
- Guzmán Riverón, Martha. 2007. *Textos del Caribe* (edition of unpublished documents from the Spanish colonial period). Doctoral dissertation. Vol. 2. Munich: Universidad Ludwig-Maximilians.
- Hidalgo Nuchera, Patricio. 2006. *Entre Castro del Río y México: correspondencia privada de Diego de la Cueva y su hermano Juan, emigrante en Indias (1601-1641)*. Córdoba: Universidad de Córdoba.
- Levisi, Margarita. 1984. *Autobiografías del siglo de oro: Jerónimo de Pasamonte, Alonso de Contreras, Miguel de Castro*, Madrid, SGEL.
- Lohmann Villena, Guillermo. 2004. Cartas de mercaderes. Secretos y confidencias en el comercio privado. *La Casa de la Contratación y la navegación entre España y las Indias*, 815-841. Sevilla, Universidad de Sevilla / CSIC.

- López Benito, Clara Isabel. 2000. La cosmovisión de una mujer salmantina emigrada a las Indias y vinculada con los Montejo de Yucatán, a través de sus cartas privadas. *Salamanca: Revista de Estudios* 44.315-367.
- Lorenzo Pinar, Francisco Javier & Vasallo Toranzo, Luis. 1996. *Diario de Antonio Moreno de la Torre. Zamora, 1673-79. Vida cotidiana en una ciudad española durante el siglo XVII*. Zamora: Instituto de Estudios Zamoranos.
- Martínez Martínez, Carmen. 2003. *Cartas y memoriales (Hernán Cortés)*. Valladolid: Junta de Castilla y León (selection of private letters).
- Martínez Martínez, Carmen. 2006. *En el nombre del hijo*. México: UNAM.
- Martínez Martínez, Carmen. 2007. *Desde la otra orilla: cartas de Indias en el Archivo de la Real Chancillería de Valladolid (siglos XVI-XVIII)*. León: Universidad de León.
- Navarro Bonilla, Diego. 2003. *Del corazón a la pluma: archivos y papeles privados femeninos en la Edad Moderna*. Salamanca: Universidad de Salamanca.
- Otte, Enrique. 1966. Cartas privadas de Puebla del siglo XVI. *Jahrbuch für Geschichte Lateinamerikas* 3.10-87.
- Otte, Enrique. 1993. *Cartas privadas de emigrantes a Indias*. Madrid: Fondo de Cultura Económica (selection of 43 letters written in the 17th century).
- Pérez Villanueva, Joaquín. 1986. *Felipe IV y Luisa Enríquez Manrique de Lara, condesa de Paredes, un epistolario inédito*. Salamanca: Caja de Ahorros de Salamanca.
- Rodríguez Morel, Genaro. 1995. Cartas privadas de Hernando Gorjón. *Anuario de estudios americanos* 52.203-233.
- Rojas, Elena. 2000. *Documentos para la historia lingüística de Hispanoamérica Vol II*. Madrid: BRAE (selection of letters written by Spaniards in the 16th and 17th centuries).
- Stoll, Eva. 1543. *La memoria de Juan Ruiz de Arce*. Madrid: Iberoamericana.
- Sánchez Rubio, Rocío, Isabel Testón Núñez & Antonio Domínguez Ortiz. 1999. *El hilo que une: las relaciones epistolares en el Viejo y en el Nuevo Mundo, siglos XVI-XVIII*. Cáceres: Universidad de Extremadura.
- Schmidt-Riese, Roland. 2002. *Relatando México. Cinco textos del período fundacional de la colonia en Tierra Firme*. Madrid, Iberoamericana.
- Szmolka Clares, José, María Amparo Moreno Trujillo & María José Osorio Pérez. 1996. *Epistolario del conde de Tendilla (1504-1506)*. Granada: Universidad de Granada (selection of letters written in 1504).

Torres, Concepción. 1995. *Ana de Jesús, cartas (1590-1621). Religiosidad y vida cotidiana en la clausura femenina del Siglo de Oro*. Salamanca: Universidad de Salamanca.

## References

- Alarcos, Emilio. 2004. *Gramática de la lengua española*. Madrid: Espasa Calpe.
- Blas Arroyo, José Luis. 2011. Deber (de) + infinitivo: ¿un caso de variación libre en español? Factores determinantes en un fenómeno de alternancia sintáctica. *Revista de Filología Española* 91:1. 9-42.
- Blas Arroyo, José Luis y Javier Vellón. 2014. La supervivencia de *deber de* + inf. en el español moderno (siglos XVIII-XX). *Revista de Filología Española* 94:1. 9-38.
- Blas Arroyo, José Luis y Margarita Porcar (2016). Patrones de variación y cambio en la sintaxis del Siglo de Oro. Un estudio variacionista de dos perífrasis modales en textos de inmediatez comunicativa. *RILCE. Revista de Filología Hispánica* 32:1. 47-81.
- Balasch, Sonia. 2008. *Debe de ser: evolución de la variación*. *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics* ed. by Maurice Westmoreland & Juan Antonio Thomas, 109-119. Somerville, MA: Cascadilla Proceedings Project.
- Balasch, Sonia. 2012. Revisión amplia de la variación *deber de* + infinitivo en la diacronía del español. *Actas del VIII Congreso Internacional de Historia de la Lengua Española* ed. by Emilio Montero, 607-620. Madrid: Arco Libros.
- Bergs, Alexander. 2005. *Social Network Analysis and Historical Sociolinguistics*. Berlin: Mouton de Gruyter.
- Bergs, Alexander. 2012. The Uniformitarian Principle and the Risk of Anachronisms in Language and Social History. *The Handbook of historical sociolinguistics* ed. by Juan Manuel Hernández-Campoy & Juan Camilo Conde Silvestre, 80-98. Malden, MA: John Wiley & Sons Inc.
- Cano, Rafael. 1996. Lenguaje ‘espontáneo’ y retórica epistolar en cartas de emigrantes españoles a Indias. *El español hablado y la cultura oral en España e Hispanoamérica* ed. by Thomas Kotschi, Wulf Oesterreicher y Klaus Zimmermann, 375-404. Madrid: Vervuert-Iberoamericana.
- Castillo Gómez, Antonio. 2006. *Entre la pluma y la pared: una historia social de la escritura en los Siglos de Oro*. Madrid: Akal.
- Conde-Silvestre, Juan C. 2007. *Sociolingüística histórica*. Madrid: Gredos.



- Davies, Mark. 2002. *Corpus del español (100 millones de palabras, siglo XIII - siglo XX)*. Available at <http://www.corpusdelespanol.org> (last search: 10-21-2013).
- Demonte, Violeta. 2001. La gramática descriptiva y la necesidad de la gramática. *El buen uso de la lengua española* ed. by Antonio Álvarez Tejedor, 92-93. Burgos: Caja de Burgos.
- Dossena, Marina. & Ingrid Tieken-Boon van Ostade, eds. 2008. *Studies in Late Modern English Correspondence: Methodology and Data*. Bern: Peter Lang.
- Elliott, John. 1998. *La España Imperial*. Barcelona: Vicens-Vives.
- Elspass, Stephan, Nils Langer, Joachim Scharloth and Wim Vandenburg, eds. 2007. *Germanic language histories 'from below' (1700-2000)*. Berlin: Walter de Gruyter.
- Fernández Álvarez, Manuel. 2004. *Sombras y luces en la España imperial*. Barcelona: Espasa.
- Fernández Alcaide, Marta. 2009. *Cartas de particulares en Indias del siglo XVI*. Madrid: Vervuert-Iberoamericana.
- Fernández de Castro, Francisco. 1999. *Las perífrasis verbales en el español actual*. Madrid: Gredos.
- García Mouton, Pilar. 1999. Tratamientos en las cartas de Indias. *Estudios de lingüística hispánica. Homenaje a María Vaquero* ed. by Amparo Morales, Eduardo Forastieri Braschi, Julia Cardona and Humberto López Morales, 263-276. San Juan: Editorial de la Universidad de Puerto Rico.
- Gómez Manzano, Pilar. 1992. Gómez Manzano, Pilar. *Perífrasis verbales con infinitivo (valores y usos en la lengua hablada)*. Madrid: UNED.
- Gómez Molina, José Ramón. 2009. Estudio sociolingüístico de la perífrasis *deber* + infinitivo y *deber de* + infinitivo en el español hablado. *Lengua viva. Estudios ofrecidos a César Hernández Alonso* ed. by Antonio Álvarez Tejedor, 867-880. Valladolid: Universidad de Valladolid.
- Hernández-Campoy, Juan Manuel & Natalie Schilling. 2012. The Application of the Quantitative Paradigm to Historical Sociolinguistics: Problems with the Generalizability Principle. *The Handbook of historical sociolinguistics* ed. by Juan Manuel Hernández-Campoy & Juan Camilo Conde Silvestre, 63-79. Malden, MA: John Wiley & Sons Inc.
- Kielkiewicz-Janowiak, Agnieszka. 2012. Class, Age, and Gender-based Patterns. *The Handbook of historical sociolinguistics* ed. by Juan Manuel Hernández-Campoy & Juan Camilo Conde Silvestre, 307-331. Malden, MA: John Wiley & Sons Inc.

- Kopaczyk, Joanna & Andreas. H. Jucker, eds. 2013. *Communities of practice in the history of English*. Amsterdam: John Benjamins.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: Pennsylvania UP.
- Labov, William. 2001. *Principles of Linguistic Change: External Factors*. Philadelphia: Blackwell Publishing Co.
- 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. 2003. *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. London: Longman Pearson Education.
- Oesterreicher, Wulf. 2004. Textos entre inmediatez y distancia comunicativas. El problema de lo hablado escrito en el Siglo de Oro. *Historia de la lengua española* ed. by Rafael Cano Aguilar, 729-769. Barcelona: Ariel.
- Ogura, Mieko & William Wang. 1996. Snowball Effect in Lexical Diffusion: The Development of –s in the Third Personal Singular Present Indicative in English. *English Historical Linguistics 1994, Current Issues in Linguistic Theory 135* ed. Derek Britton, 119-141. Amsterdam: John Benjamins.
- Okulska, Urszula. 2010. Performing the world of politics through the discourse of institutional correspondence in Late Middle and Early Modern England. *Perspectives in Politics and Discourse* ed. by In Urszula Okulska & Piort Cap, 173-97. Amsterdam: John Benjamins.
- Pahta, Pälvi. & Andreas. H. Jucker, eds. 2011. *Communicating early English manuscripts*. Cambridge: Cambridge University Press.
- Penny, Ralph. 2000. *Variation and Change in Spanish*. Cambridge: Cambridge University Press.
- Raumolin-Brunberg, Helena. 2005. Language change in adulthood: Historical letters as evidence. *European Journal of English Studies* 9:1. 37-51.
- Romaine, Suzanne. 1982. *Socio-Historical Linguistics*. Cambridge: Cambridge University Press.
- Tagliamonte, Sali. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley-Blackwell Publishers.
- Tagliamonte, Sali A. & Jennifer Smith. 2006. Layering, change and a twist offate: Deontic modality in dialects of English. *Diachronica* 23:3. 341-80.

Włodarczyk, Matylda. 2013. Thematic issue on historical sociolinguistics and pragmatics of migrant contexts. *Poznan Studies in Contemporary Linguistics* 49:3. 275-279.

Yllera, Alicia. 1980. *Sintaxis histórica del verbo español: Las perífrasis medievales*. Zaragoza: Universidad de Zaragoza.