

Mariana Gomes\* ([marianaferreiragomes@gmail.com](mailto:marianaferreiragomes@gmail.com))  
Leonor Tavares\* ([lactavares@yahoo.com.br](mailto:lactavares@yahoo.com.br))  
Ana Rita Guilherme\* ([anabrugui@gmail.com](mailto:anabrugui@gmail.com))

\*University of Lisbon, Linguistics Center of the University of Lisbon (CLUL)  
([www.clul.ul.pt](http://www.clul.ul.pt))  
Phone Number: 351 21 790 47 00

“*estas minhas limitadas cifras tenham a felicidade de acharem a VMce. desfrutando aquela saúde espiritual e corporal tão feliz como lhe deseja o meu afecto*” - Different perspectives on correspondance conventionalities[1]

### Resumen

Este artículo presenta nuestra actual experiencia de edición crítica electrónica de cartas privadas portuguesas involucrando mas de 2000 hombres e mujeres de todos los estratos sociales y que han vivido en la época Moderna (siglos XVI a XIX). Las cartas son tratadas como fuentes para varias perspectivas intelectuales, incluyendo la Pragmatica Historica (Jacobs and Jucker 1995), la perspectiva más relevante en este artículo. El objetivo último de nuestro proyecto, sin embargo, es articular metodologias diferentes en la análisis de textos intentando evocar "murmuring voices of societies" ("voces murmurantes de las sociedades"), como Michel de Certau a llamado al comportamiento creativo de los individuos en su vida cotidiana.

### Abstract

This paper arises from our current experience of making the critical electronic edition of Portuguese private letters involving more than 2.000 men and women from all social strata who lived in the Modern era (16<sup>th</sup> to 19<sup>th</sup> century). The letters are being treated as sources for various intellectual approaches, including Historical Pragmatics (Jacobs and Jucker 1995), the one approach that is relevant for this paper. The final aim of the project, though, is to articulate different methodologies in analysing texts in an effort of conjuring the old “murmuring voices of societies”, as Michel de Certeau called the creative behaviour of individuals in everyday life.

#### 1. Sources for the study of Historical Pragmatics

Historical Pragmatics has been attracting those scholars who show an interest on the situated uses of language in history; as Fitzmaurice and Taavitsainen put it, “[it] offers insights into earlier communicative practices, registers, and linguistic functions as gleaned from historical discourse” (Fizmaurice and Taavitsainen 2007:1). The discipline follows, very naturally, the current trends in linguistic inquiry, so it has adopted the same cautious method of dealing with large quantities of data – i.e. computer-based corpora – since this posture manifestly allows for linguistic analyses to present themselves as empirically reliable (Taavitsainen and Fitzmaurice 2007:16–17).

However, as all pragmaticists promptly recognize, a linguistic *corpus* with encoded contextual information – hence language usage information – is no easy object to obtain (Kohnen 2007). In the depths of a *corpus* marked-up for parts of speech, for instance, the unity of the texts to which language utterances once belonged becomes totally lost. Lost is also, in this way, the meaning language usage got from the discursive and extra-linguistic surroundings it belonged to.

So we should ask: – When was it that the historical study of language had a systematic way of accounting for texts in context? Answer: – When it was called Philology and didn’t diverge from literary studies or textual criticism. Now, the subsequent question would be: – How did those old partners react to computerized technology and the subsequent facility of accounting for large quantities of data? – Well, literary studies can rely today on their own electronic *corpora*, i.e. hyper

textual editions of literary texts prepared by textual critics who gained computer science competences (cf., for instance, Robinson 1997). It becomes clear that the methodology used today on digital critical editions should be useful, in principle, to Historical Pragmatics research.

Reasoning like this, we prepared a diachronic *corpus* of Portuguese letters following conventions currently used within Textual Criticism to produce digital literary editions: these conventions belong to the well-known project TEI, *Text Encoding Initiative*, and to a variant of it, DALF, *Digital Archive of Letters in Flanders*, especially conceived for the letters' genre critical edition.

## 2. The corpus for this study

Our sources were carefully chosen as relevant sources from the point of view of Historical Pragmatics: the speakers who wrote them came from several social backgrounds and the letters were written on several different local contexts. That kind of complete representativeness was firstly made possible because the manuscripts we use were largely kept within judicial court-files from the Portuguese Inquisition (1536-1822) and the Portuguese Royal Appeal Court (late 18<sup>th</sup> century-1836), surrounded with documented testimonies on the accused individuals who wrote or received the written correspondence.

The information on social backgrounds and local contexts, being thus partially recoverable, gets also encoded in the electronic critical edition. By following Textual Criticism methodologies while keeping in mind social analyses preoccupations, we manage to connect such diverse information as manuscript's physical layout, original writing, authorial emendations, editorial conjectures, information on variants (when copies or comments also survived), information on the context of the letter's composition (an event within a social context), and information on the participants' biographies (cf. Table 1).

**Table 1**

### Abbreviated description of the *corpus*

The *corpus* includes the following main fields (with several sub-divisions):

- Letter title
- Sponsor and Funder
- Project Identification
- Archival Identification
- Letter's Identifier (author, receiver, scribe, annotator, time and place)
- Letter's synopsis
- Letter's context
- Letter's transcription

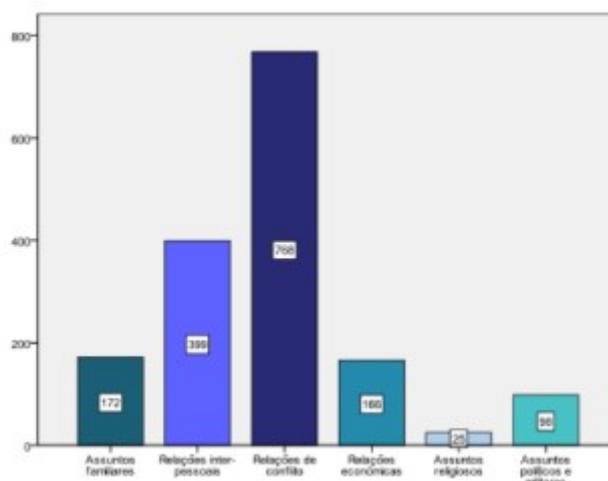
The manuscripts' transcription is quasi-diplomatic, allowing only for regular word separation. Authorial strikethroughs and additions, line-breaks, lacunae, difficult deciphering, abbreviations and non-orthographic uses are respected in the transcription by means of XML tags specifically developed by TEI to fit primary sources idiosyncrasies. External to this electronic support, but linked to it by anchors, there is a database containing biographic information on authors and addressees.

As for the corpus dimensions, it contains, for now, 450.000 words, more than 1.600 letters and information on more than 2.000 participants (several letters were written or received by the same person).

Of course the themes of letters apprehended by court officials, as can be seen on Table 2, were different, in terms of frequency, from the ones figuring in the total number of letters that actually circulated in those times: letters related to conflict relations, for instance, are much more frequent in our *corpus* than the banal example of letters on family affairs:

**Table 2**

### Thematic distribution of letters within the *corpus*



172 – Family affairs                      399 – Interpersonal relations                      768 – Conflict relations  
 166 – Economic relations                      25 – Religion affairs                      98 – Politics and military affairs  
 Total – 1.628 (these results go back to October 2009)

On the other hand, the sample has a large geographical representation, along with a fairly good chronological one (Table 3). The Royal Appeal Court (*Casa da Suplicação*) judged lawsuits from all over the reign and from the overseas empire; the Lisbon Inquisition covered the reign’s central area and the overseas territories; the Coimbra Inquisition covered the reign’s north, and the Évora Inquisition covered the south. We used also some few lawsuits from the late XIXth century coming from the north-eastern part of the reign (Bragança):

**Table 3**  
**Areal and chronological distribution of the transcribed letters**



### 3. An analysis of formulaic excerpts within the letters

One of the most important groups of tags we use in our *corpus*, considering the Historical Pragmatic point of view, is the protocol/eschatocol group. Letters can be highly conventionalized in matters of textual openings and closings. Although the ones from the Modern period didn’t keep exactly the mediaeval structure codified in the *ars dictaminis* tradition (Guillén 1986), one can still find a large amount of fossilized expressions surrounding the expository text in those 16<sup>th</sup> to 19<sup>th</sup> century letters. This means that for matters of automatic search, looking for linguistic structures, for instance, within the letters’ text, it is important to hide either the formulaic part of the *corpus*, or its expository one.

The system we adopted was partly dictated by DALF conventions, which allow for a mark-up of the external part (in visual terms) of letters – date, place, addressing formula – and partly dictated by TEI conventions, which has a tag for text segments that can receive an arbitrary attribute; the

chosen attributes for the segments were precisely “hargue”, for the formulaic introduction, and “peroration”, for the formulaic conclusion within the letters’ visual body. An example goes like this (Tables 4 and 5):

**Table 4**

**Protocol (external opener+hargue) from a letter sent to her mother by a Brazilian farmer’s daughter, 18<sup>th</sup> century**

```
<opener>
<salute>Minha Mãe <abbr>Sr. ra</abbr></salute>
</opener>
<p><seg type="formulaicText" n="hargue">estimarei que
estas limitadas regras<lb/>a axhem a <abbr>vme</abbr>
assistida di perfeita saude em companhia<lb/>e Meu Pai e
<abbr>Snro :/</abbr> e de toda a nobre ciza para disporem da
minha<lb/> que o prezente he boa seja Deus
louvado<lb/></seg>
```

*My mother and lady*

*I hope these few lines will find you in perfect health in the company of my father and lord and of all the noble house, so that you can dispose of mine that is good at present, praise the Lord*

**Table 5**

**Eschatocol (peroration+external closer) from a letter sent from a Portuguese architect to a mathematician, 18<sup>th</sup> century**

```
<seg type="formulaicText" n="peroration">que Deus Guarde
por felices annos <seg></p>
<closer>
<salute>De <abbr>VMce</abbr><lb/> Muito attento Venerador, e
Verdadeiro amigo</salute>
<signed>Mathias Francisco de Pazos</signed>
</closer>
```

*... may the Lord guard (Your Mercy) for happy years  
The venerating attentive and true friend of Your Mercy  
Mathias Francisco de Passos*

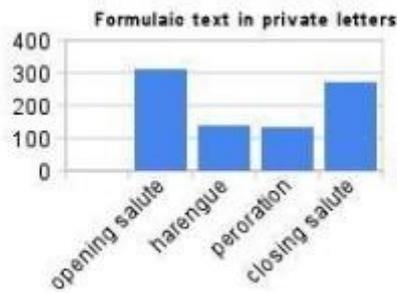
This sort of mark-up allows for an automatic distinction between letters presenting formulaic utterances, along with letters missing them. We have analyzed a sample of 182.000 words, 362 letters, from the late 16<sup>th</sup> (11), 17<sup>th</sup> (100), 18<sup>th</sup> (188) and early 19<sup>th</sup> century (63). Besides from the chronological one, the sample has also some balanced representativeness in terms of gender (Table 6). Male authors (268) and receivers (281) are more present than female ones, but the feminine universe is nevertheless there, with 94 women-authors and 44 women-receivers.

**Table 6**



Looking for the repetitive expressions in their beginnings and endings, we arrived at the following results (Table 7): 449 protocols vs 423 eschatocols; 309 opening salutations vs 289 closing salutations; 138 hargues vs 132 perorations.

**Table 7**



So far, we could conclude that, in a socially balanced sample of private letters written in Early Modern and Late Modern Portuguese history, i) the presence of formulaic utterances is the rule, while its absence is the exception; ii) opening formulaic utterances are more frequent than closing ones; iii) given the letter body, external pieces of those repetitive expressions (salutations) are more frequent than the internal ones (harengues and perorations). A cross-examination of the authors' biographies gives a good indication of people from low social strata apparently haranguing and perorating a lot in letters, while *élite* letter writers largely dispensed the strategy. (Precise numbers on this point will be given at the Conference).

Although it goes without saying that we are dealing with a narrow sample, some temporary interpretations can be made regarding the formulaic behaviour of our letter writers.

On the one hand, physical factors must not be discarded: the blank area while beginning a letter is forcefully wider than the available one when ending it; practical writers must put up with that, finishing the letter inside the available space, and thinking no more of it. But besides that chance factor pending on spontaneous reactions, those letter writers were also social actors engaged in verbal and non verbal deferred interaction, so the formulaic parts in their texts must also have been partly meaningful to them. This means that the semantic fields covered by the utterances from the protocols and eschatocols in the sample must also have been relevant in the pragmatic behaviours of those time.

Semantic classification of historical texts can be very biased, so we followed the same choice made by Archer and Culpeper (2009) in their study of language usage in a close-to-spoken *corpus* similar to our own, namely dialogues in plays and trial proceedings of the 17<sup>th</sup> and 18<sup>th</sup> centuries; this means we also applied the *UCREL Semantic Analysis System* (USAS), and selected from it the following relevant categories:

<b>Semantic Group</b>	<b>Sub-classification</b>
<b>B:</b> The Body and the individual	<b>B2:</b> Health and diseases
<b>E:</b> Emotional actions, states and processes	<b>E2:</b> Liking
<b>Q:</b> Linguistic actions, states and processes	<b>Q1.1.:</b> Communication in general
<b>S:</b> Social actions, states and processes	<b>S1.2.4.:</b> Politeness
	<b>S3.1.:</b> Relationship in general
	<b>S4:</b> Kin
	<b>S7.2.:</b> Respect
	<b>S9:</b> Religion

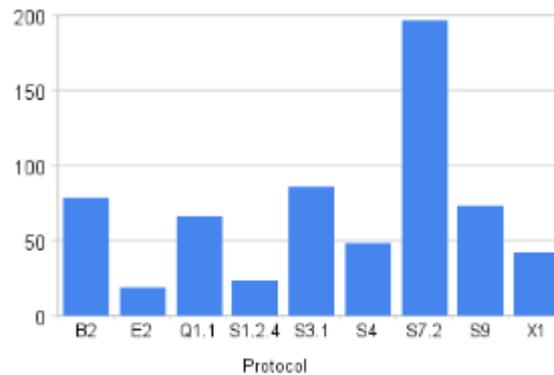
X: Psychological actions, states and processes

X1: General

The presence of these semantic fields in the sample's formulaic utterances is summed-up in tables 9 and 10:

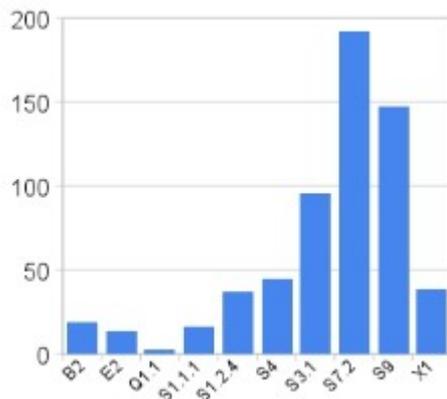
**Table 9**

**Semantic fields in private letters' protocols**



**Table 10**

**Semantic fields in private letters' eschatocols**



Description:

(i) In the textual segment protocol, the semantic field of "respect" stands out - it appears 196 times - followed by terms related with general relations such as friendship - 86 times. The "health" semantic field is also very prominent - it appears 79 times in the harangue segment-, as does the "religious" semantic world - 73 times.

(ii) Other semantic fields found in the protocol are: "communication in general" (66); "kin" (48); general expressions related with "psychological actions, states and processes" (42); "politeness" (23); and "liking emotions" (19).

(iii) All the above mentioned semantic fields were also found in the closing formulaic textual segment- the eschatocol. And, similar to what happened in protocols, the semantic field of "respect" is the most expressive - it appears 192 times -, followed by the "religion" semantics - 147 times. Reference to "health" is less expressive than in the harangue - only 19 times. Also not very expressive are expressions concerning "communication in general" (3 times).

(iv) The other semantic fields in the eschatocol have the following distribution: terms related with general relations such as friendship - 96 times; "kin" words - 45 times; references to "psychological actions, states and processes" appears 39 times; "politeness" expressions such as "não enfado mais"

(*I won't bother you longer*) appear 37 times in the *corpus*; "goodbye" expressions as "adeus, adeus" were counted 16 times and "liking emotions", 14 times.

(v) Taking into account all the semantic fields observed in the formulaic segments of the letters, the most expressive group is the one concerned with Social Actions, States and Processes: the most prominent semantic fields found in our *corpus*, such as "respect", "religion", "kin" words and terms expressing general relations fall under the scope of that larger group.

Although this is a strictly descriptive study, we think this new kind of data can be very informative to Historical Pragmatics in terms of textual linguistics (Adam 1992/2005), politeness in language (Brown and Levinson 1987) and discourse analysis (Fairclough 2003, Van Dijk 1997). We will bring to the Conference some final remarks on those theoretical implications.

## References

- Adam, Jean-Michel (1992/2005) *Les textes: types et prototypes. Récit, description, argumentation, explication et dialogue*. Paris: Nathan.
- Archer, Dawn Archer and Jonathan Culpeper (2009) "Identifying key sociophilological usage in plays and trial proceedings (1640–1760): an empirical approach via corpus annotation", *Journal of Historical Pragmatics*, 10:2, pp. 286–309.
- Brown, Penélope and Stephen C. Levinson (1987) *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press (2<sup>nd</sup> ed.)
- Certeau, Michel (1990/2002) *L'Invention du quotidien, 1. Arts de faire*. Paris: Gallimard.
- DALF, <http://www.kantl.be/ctb/project/dalf/dalfnl.htm> (accessed Feb. 15, 2010).
- Fairclough, Norman (2003): *Analysing Discourse: Textual Analysis for Social Research*. Londres: Routledge.
- Fitzmaurice, Susan, and Irma Taavitsainen (2007) "Introduction", in Susan M. Fitzmaurice and Irma Taavitsainen (eds.), *Methods in Historical Pragmatics*. Berlin/New York: Mouton de Gruyter, pp. 1–10.
- Guillén, Claudio (1986) "Notes towards the study of the Renaissance letter", in Barbara Kiefer Lewalski (ed.), *Renaissance Genres: Essays on Theory, History and Interpretation*. Cambridge, MA: Harvard University Press, pp. 70–101.
- Jacobs, Andreas, and Andreas Jucker (1995) "The historical perspective in pragmatics", in Andreas H. Jucker (ed.), *Historical Pragmatics: Pragmatic Developments in the History of English*. Amsterdam/Philadelphia: John Benjamins, pp. 3–33.
- Kohonen, Thomas (2007) "Text types and the methodology of diachronic speech act analysis", in Susan M. Fitzmaurice and Irma Taavitsainen (eds.), *Methods in Historical Pragmatics*. Berlin/New York: Mouton de Gruyter, pp. 139–166.
- Robinson, Peter, 1997, "A stemmatic analysis of the fifteenth-century witness to *The Wife of Bath's prologue*", in Blake and Robinson (eds.), *The Canterbury Tales Project's Occasional Papers*. Oxford: Office For Humanities Communication, pp. 69–132.
- Taavitsainen, Irma, and Susan Fitzmaurice (2007) "Historical pragmatics: what is it and how to do it?", in Susan M. Fitzmaurice and Irma Taavitsainen (eds.), *Methods in Historical Pragmatics*. Berlin/New York: Mouton de Gruyter, pp. 11–36.
- TEI, <http://www.tei-c.org/index.xml> (accessed Feb. 15, 2010).
- UCREL Semantic Analysis Systems, <http://ucrel.lancs.ac.uk/wmatrix2.html> (accessed Feb. 15, 2010).
- Van Dijk, Teun A. (1997) "The study of discourse", in Teun A. van Dijk (ed.), *Discourse as Structure and Process*. Londres: Sage, pp. 1–34.

---

[1] Paper written with the collaboration of Rita Marquilhas (CLUL- University of Lisbon, Linguistics Center of the University of Lisbon)