

USERGUIDE FOR DIGITAL EDITION OF TEXTS IN *P. S. POST SCRIPTUM*

Gael Vaamonde

English version by Clara Pinto

Last updated: 11/01/2017

Table of contents

0. INTRODUCTION-----	5
1. THE HEADER: <TEIHEADER> -----	7
1.1. <fileDesc>-----	7
1.1.1. <titleStmt>-----	7
1.1.1.1. <title>-----	7
1.1.1.2. <sponsor>-----	11
1.1.1.3. <funder>-----	11
1.1.1.4. <respStmt>-----	12
1.1.1.5. <Example>-----	13
1.1.2. <publicationStmt>-----	14
1.1.2.1. Example -----	14
1.1.3. <sourceDesc>-----	15
1.1.3.1. <letIdentifier>-----	15
1.1.3.1.1. <institution>-----	16
1.1.3.1.2. <repository>-----	16
1.1.3.1.3. <idno>-----	16
1.1.3.1.4. <altName>-----	17
1.1.3.1.5. Example -----	17
1.1.3.2. <letHeading>-----	18
1.1.3.2.1. <author> and <addressee>-----	18
1.1.3.2.2. <placeLet>and <placeDest>-----	21
1.1.3.2.3. <dateLet> and <date>-----	24
1.1.3.2.4. Example -----	24
1.1.3.3. <physDesc>-----	26
1.1.3.3.1. <type>-----	27
1.1.3.3.2. <support>-----	27
1.1.3.3.3. <extent>-----	29
1.1.3.3.4. <layout>-----	30
1.1.3.3.5. <decoration>-----	34
1.1.3.3.6. <condition>-----	37
1.1.3.3.7. Example -----	37
1.1.3.4. <envOcc>-----	38
1.1.3.5. <letContents>-----	39
1.1.3.5.1. <locus>-----	39
1.1.3.5.2. <class>-----	40
1.1.3.5.3. Summary of the letter-----	42
1.1.3.5.4. Example -----	42
1.1.3.6. <additional>-----	45
1.1.3.6.1. <adminInfor>-----	45
1.1.3.6.2. <surrogates>-----	45
1.1.3.6.3. Letter context -----	46
1.1.3.6.4. Translation to English -----	47
1.1.3.6.5. Example -----	48
1.2. <encodingDesc> -----	48
1.3. <profileDesc> -----	49

1.4. <revisionDesc>	49
1.5. Full example of <teiHeader>	52
2. THE TEXT: <TEXT>	56
2.1. <envelope>	56
2.2. <body>	56
2.2.1. <opener> and <closer>	57
2.2.2. <ps>	62
2.3. Full example of <text>	63
3. TRANSCRIPTION AND EDITING CONVENTIONS	64
3.1. Standardization of punctuation	64
3.2. Annotations appended to the text	64
3.3. Abbreviations	65
3.4. Beginning of folio <pb @n/>	68
3.5. Beginning and end of paragraph <p></p>	70
3.6. Beginning of line: <lb/>	70
3.7. Vertical text: <lb @line_rend/>	71
3.8. Corrected text: 	72
3.9. Added text: <add>	73
3.10. Underlined text: <hi>	74
3.11. Omitted text: <gap/>	75
3.12. Supplied text: <supplied>	77
3.13. Unclear text (<i>locus desperatus</i>): <unclear>	79
3.14. Text from another hand: <handShift/>	79
3.15. Discourse organization: <s formula>	80
3.16. Postmark: <postmark>	81
3.17. Other cases:	81
4. XML CHANGES IN TEITOK	83
4.1. General explanation	83
4.2. Example	85
5. CONVERSION TO TEI-P5	88
5.1. General explanation	88
5.2. Internal model	88
5.3. External model (TEI-P5)	91
5.4. Main differences between the two models	96
5.4.1. Editing tasks	96
5.4.2. Communicative framework	96
5.4.3. Summary of the letter	97
5.4.4. Context of the letter	98
5.4.5. Translation of the letter	98
5.4.6. Image description	99
5.4.7. Content classification	99
5.4.8. Facsimile images	100
5.4.9. Envelope	101
5.4.10. Postscript	101

5.4.11.	Line breaks -----	102
5.4.12.	Redundant or superfluous text: <surplus> -----	102
5.4.13.	Others -----	102
5.4.14.	Attachment text encoding -----	104
6.	THE BIOGRAPHIC DATABASE -----	107
6.1.	The header -----	107
6.2.	The participants list-----	107
6.2.1.	<person> -----	109
6.2.2.	<persName>-----	109
6.2.3.	<affiliation>-----	110
6.2.4.	<birth> and <death>-----	110
6.2.5.	<education>-----	111
6.2.6.	<event>-----	111
6.2.7.	<faith>-----	112
6.2.8.	<floruit> -----	112
6.2.9.	<langKnowledge>-----	112
6.2.10.	<nationality> and <residence>-----	112
6.2.11.	<socecStatus>-----	115
6.2.12.	<state>-----	116
6.2.13.	<trait>-----	116
6.3.	Example of a biographic record -----	117

0. Introduction

This manual presents an explanation of the criteria adopted in the encoding of texts using XML markup, within the Project *P.S. Post Scriptum*. Similarly to current practices in the field of Digital Humanities, *P.S. Post Scriptum* follows the standards proposed by the TEI (*Text Encoding Initiative*) consortium for text edition in digital format. However, three remarks, motivated by the project's own methodology, should be made.

1. For the XML encoding of the texts we adopt the model proposed by the Project DALF (*Digital Archive of Letters by Flemish Authors and Composers from the 19th & 20th century*), which was itself based on the P4 version of the TEI consortium. Nonetheless, adapting this model to *P.S. Post Scriptum* required certain changes, which were imposed by the nature of the texts, but also by the double purpose of *P.S. Post Scriptum*: to offer a scholarly digital edition alongside a linguistically annotated corpus. The result is a customized model, whose DTD is only valid within the *Project P.S. Post Scriptum* (dalf-tt.dtd).
2. This internal model is prepared to include linguistic information and purely philological information. The linguistic treatment of data offered by *P.S. Post Scriptum* is done using the TEITOK online environment, developed by Maarten Janssen. In this environment texts are tokenized, spelling is normalized and lemmatization and morphosyntactic annotation are generated.
3. For users interested in a standardized model, we make available a TEI-P5 conformant version largely based on the choices made by the CHARTA (*Corpus Hispánico y Americano en la Red: Textos Antiguos*¹) network.

These guidelines were mainly conceived to work as a reference manual for internal use of the *P.S. Post Scriptum* team. Therefore, most references made here bear in mind the customized model (presented in 1.), which we will, from now on, refer to as the “internal model”.

We will, however, dedicate a brief section to the explanation of the XML elements that are added to the electronic edition when it enters the TEITOK environment (cf. section 4). As for the correspondences with the TEI-P5 standard, they are pointed out, one by one, throughout the manual. They are also summarized at the end, where we include a section illustrating such model, which we will call “external model” (cf. section 5).²

¹ Available at <files.charta.es/200000059-d00b8d104d/Guía%20CHARTA-TEI.pdf>

² Users interested in downloading any of these two models should bear in mind that what is called here “internal model” and “external model” are referred to as on the website as “TEITOK XML” and “Pure TEI P5 XML”, respectively.

For the sake of clarity, all the information regarding paragraphs 2 and 3 (i.e. linguistic treatment with TEITOK and standardization of the XML model TEI-P5) will appear with a grey background.



The main purpose is to distinguish the information intended to guide the transcription tasks, which is the main goal of this users guide, from the rest of the information, which goes beyond transcription tasks.

1. The Header: <teiHeader>

The element <teiHeader> contains four main parts:

- <fileDesc> contains complete bibliographic description of the digital file.
- <encodingDesc> contains information concerning the relation between digital edition and original source.
- <profileDesc> contains non-bibliographic information concerning the text: language, context, participants, etc.
- <revisionDesc> contains the history of revisions made to the digital file.

1.1. <fileDesc>

This element contains the full bibliographic description of the digital file and includes the following elements:

- <titleStmt> contains information about the title of the letter and the identification of the people responsible for the different edition tasks.
- <publicationStmt> contains information about the host institution and the Project responsible for publishing the digital edition.
- <sourceDesc> contains information about the source which is the base of the digital edition.

1.1.1. <titleStmt>

This element contains the following parts:

- <title> contains the title of the digital file.
- <sponsor> contains the name of the host institution.
- <funder> contains the name of the institution funding the project.
- <respStmt> contains the name of the people responsible for different aspects of the edition.

1.1.1.1. <title>

- a. The content of the element <title> always obeys to the following structure:

```
<title>Año. Carta de X para Y.</title>
```

Example:

```
<title>1646. Carta de Miguel Gutiérrez Peña para Felipe Díaz.</title>
```

- b. Optionally, the author's or the addressee's occupation, their nobility titles and/or the kinship relation between them may be added:

Example 1:

<title>1644. Carta de Pedro de Baltanas para su sobrina Lorenza de Baltanas.</title>

Example 2:

<title>1644. Carta de doña María Antonia Godoy, condesa de Valdegrana, para don Mariano de los Ríos Fernández de Córdoba, conde de Gavia.</title>

Example 3:

<title>1823. Carta de Rodrigo José dos Santos, ourives, para Diogo Jacinto de Almeida, escrivão.</title>

- c. Conjectured elements are encoded inside square brackets:

Example 1:

<title>[1707]. Carta de Manuel de Santa Leocadia para su hijo de confesión Pedro Pablo Díez, boticario.</title>

Example 2:

<title>1642. Carta de João Carreira, capelão, para [a Inquisição de Lisboa].</title>

If it is not possible to conjecture a year, an interval should be conjectured:

Example 3:

<title>[1826-1828]. Carta de Manuel Fernández, coronel comandante general, para Manuel López, capitán de caballería.</title>

- d. If the letter is not an autograph (if it was dictated by a mental author whose name appears in the signature) the procedure should be the following:

Example from PS6245

<title>1710. Carta no autógrafo de María Santos de Olalla para Santiago Caballero, alguacil de campo.</title>

Furthermore, the existence of a scribe must be declared inside <letHeading> (cf. 1.1.3.2.1.g.):

<respStmt><resp subcat="scribe" id="scribe"/></respStmt>

The corresponding biographic file is always that of the mental author of the letter and not that of the scribe (cf. 6.2.2.e.)

- e. If the letter was dictated to a scribe and signed by the scribe, the procedure should be the following:

<title>Año. Carta escrita por X en nombre de Y para Z.</title>

X: name of the scribe

Y: name of the mental author

Z: name of the addressee

Example:

Example from PSCR6642:

<title>[1750-1760]. Carta escrita por Salvador de Barrionuevo en nombre de Josefa del Valle, tendera, para Fernando del Campo, velero.</title>

Furthermore, a scribe must be declared inside <letHeading> (cf. 1.1.3.2.1.g.):

<respStmt><resp subcat="scribe" id="scribe"/></respStmt>

The corresponding biographic file is always that of the mental author of the letter and not that of the scribe (cf. 6.2.2.e.).

- f. If the letter is apocryphal (a letter deceitfully written in the name of someone else) the procedure should be the following:

<title>Año. Carta apócrifa de X fingiendo ser Y para Z.</title>

X: name of the real author

Y: name of the person whose name is in the signature

Z: name of the addressee

Example:

Example from PSCR1409:

<title>1634. Carta apócrifa de Manuel da Costa fingindo ser Francisco Simões, marinheiro de carreira da Índia, para Domingos Francisco (o Chapado), barbeiro de espadas.</title>

The biographic file is always that of the real author of the letter and not that of the fake author.

- g. If the letter was copied from a draft (the hand of the letter and the hand of the draft are not the same) the procedure should be the following:

<title>Año. Carta de X, copiada del borrador de Y, para Z.</title>

X: name of the autor of the letter

Y: name of the autor or the draft

Z: name of the addressee

- h. If one of the participants uses a fake name (e.g. José Sáez Cantos) but the court proceeding allows us to know his or her real name (e.g. José Cantos Villasante), the procedure should be the following:

Example from PS6142:

<title>1758. Carta de Casimiro González, cura, para José Sáez Cantos, pseudónimo de José Cantos Villasante, sacristán y maestro de primeras letras.</title>

Only the real name is declared inside <letHeading>:

```
<addressee attested="yes" key="CDD.xml#JCV1">José Cantos Villasante</addressee>
```

In the corresponding biographic file, the element **<name>** must be filled with the real name, not the false one.

```
<persName><name>José Cantos Villasante</name></persName>
```

In the corresponding biographic file, the element **<desc>** must have a note about the use of a fake name.

```
<event><desc>usaba el pseudónimo de José Sáez Cantos</desc></event>
```

- i. If there are two authors with different kinship relations regarding the addressee, the procedure should be the following:

Example from PSCR5253:

```
<title>1606. Carta de María Vaz y de Estêvão Gonçalves para Pedro Gonçalves, marido de la primera y hermano del segundo.</title>
```

- j. If the author (X) and/or the addressee (Y) is a non-identified participant, the procedure should be the following:

Example from PS4036:

```
<title>[1690-1700]. Carta de autor no identificado para Diego Blanco.</title>
```

Example from PS6034:

```
<title>[1829]. Carta de José Moro para un destinatario no identificado.</title>
```

Example from PS6041:

```
<title>181. Carta entre participantes no identificados.</title>
```

For more information on how to treat unidentified participants inside element **<letHeading>**, see 1.1.3.2.1.d.

- k. If the author is anonymous, the procedure should be the following:³

Example from PS5019:

```
<title>Año. Carta anónima para Gabriel Lorenzo Toribio y su yerno Joaquín de Molina y Torres.</title>
```

- I. If the author is only known by his *alias*, the procedure should be the following:

```
<title>Año. Carta de autor no identificado, de apodo El Terrorista, para Y.</title>
```

- m. If the author's name and *alias* are known, the procedure should be the following:⁴

³ The difference between a non-identified participant and an anonymous participant relies on the intentionality. In the first case, the identity is unknown due to lack of data, while in the second case the author intentionally hides his identity (usually in threat or accusation letters). For more information on procedures in both cases within the element **<letHeading>**, see sections 1.1.3.2.1.d and 1.1.3.2.1.e.

Example from PSCR6130:

<title>1599. Carta de Alonso de Madrid (el Joven), tendero, para su padre Alonso de Madrid (el Viejo).</title>

Inside <letHeading> only the name is declared, not the *alias*.

```
<author attested="yes" key="CDD.xml#AM22" id="AM22">Alonso de Madrid</author>
<addressee attested="yes" key="CDD.xml#AM23">Alonso de Madrid</addressee>
```

Inside the element <name>, the existence of the real name and the *alias* is signaled as below:

```
<persName><name>Alonso de Madrid, alias el Joven</name></persName>
```

- n. If the letter is a copy and not an original, the procedure should be the following:

Example from PS6021:

<title>1705. Copia de carta de Juana de Arriola para su hijo Francisco del Puerto, marinero.</title>

Furthermore, the keyword “copy” should be added in the element <class subcat="linguisticSource"> (cf. 1.1.3.5.2.c). Copies are transcribed and modernized, but they don't receive POS, syntactic annotation or linguistic keywords (Itags), since they are not part of the linguistic corpus of the P.S. Post Scriptum project.

- o. If the letter is a draft, the procedure is the following:

```
<title>Año. Borrador de carta de X para Y.</title>
```

1.1.1.2. <sponsor>

This element contains the name of the host institution that publishes the corpus.

```
<sponsor>CLUL, Centro de Linguística da Universidade de Lisboa</sponsor>
```

1.1.1.3. <funder>

This element contains the name of the institution that funds the project. Its content is invariable.

```
<funder>ERC, European Research Council - 7FP/ERC Advanced Grant 2011
GA295562</funder>
```

⁴ On the treatment of participants with an *alias* in the biographic data record, see section 6.2.2.c.

1.1.1.4. *<respStmt>*

This element includes the list of people responsible for the different tasks in the edition of the text. It contains an element **<resp @n>** to be filled with the names of the people responsible for a certain task, as well as an element **<name @id>** to be filled with the initials and the name of each person. Each group of initials can only be declared once.

```
<respStmt>
    <resp subcat="transcription"><name id="GV">Gael Vaamonde</name></resp>
</respStmt>
```

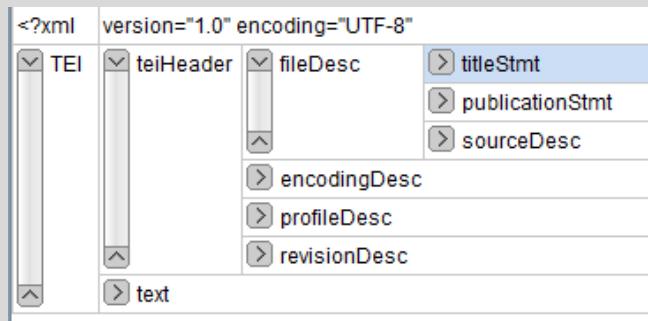
1.1.1.5. <Example>

Example from letter PS4001:

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="CardsFly1.xsl"?>
<!DOCTYPE TEI.2 PUBLIC "-//CTB//DTD Dalf 1.0 (based on TEI)//NL" "dalf-tt.dtd">
<TEI.2>
<teiHeader>
<fileDesc>
<titleStmt>
    <title>[1542]. Carta de Pedro López de la Cañada, tañedor de vihuela, para su hermano Diego Martínez de la Cañada, clérigo.</title>
    <sponsor>CLUL, Centro de Linguística da Universidade de Lisboa</sponsor>
    <funder>ERC, European Research Council - 7FP/ERC Advanced Grant 2011 GA295562</funder>
    <respStmt><resp subcat="project"><name id="PS">PS, Post Scriptum</name></resp></respStmt>
    <respStmt><resp subcat="coordination"><name id="RM">Rita Marquillas</name></resp></respStmt>
    <respStmt><resp subcat="transcription"><name id="CSS">Carmen Serrano Sánchez</name></resp></respStmt>
    <respStmt><resp subcat="posannotation"><name/></resp></respStmt>
    <respStmt><resp subcat="parsing"><name/></resp></respStmt>
    <respStmt><resp subcat="revision"><name/></name></resp></respStmt>
    <respStmt><resp subcat="modernizedEdition"><name id="GV">Gael Vaamonde</name></resp></respStmt>
    <respStmt><resp subcat="keywords"><name/></resp></respStmt>
    <respStmt><resp subcat="contextualization"><name>Carmen Serrano Sánchez</name></resp></respStmt>
</titleStmt>
```

TEI-P5 Correspondence with the example 1.1.1.5.

Structural position of <titleStmt>:



Encoding of the content in <titleStmt>:

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:id="ES">
<teiHeader>
<fileDesc>
<titleStmt>
    <title>[1542]. Carta de Pedro López de la Cañada, tañedor de vihuela, para su hermano Diego Martínez de la Cañada, clérigo.</title>
    <sponsor>CLUL, Centro de Linguística da Universidade de Lisboa</sponsor>
    <funder>ERC, European Research Council - 7FP/ERC Advanced Grant 2011 GA295562</funder>
    <respStmt><resp>Proyecto</resp><name xml:id="PS">PS, Post Scriptum</name></resp></respStmt>
    <respStmt><resp>Coordinación</resp><name xml:id="RM">Rita Marquillas</name></resp></respStmt>
    <respStmt><resp>Transcripción</resp><name xml:id="CSS">Carmen Serrano Sánchez</name></resp></respStmt>
    <respStmt><resp>Anotación morfosintáctica</resp><name/></resp></respStmt>
    <respStmt><resp>Anotación sintáctica</resp><name/></resp></respStmt>
    <respStmt><resp>Edición modernizada</resp><name xml:id="GV">Gael Vaamonde</name></resp></respStmt>
    <respStmt><resp>Revisión</resp><name/></resp></respStmt>
    <respStmt><resp>Traducción</resp><name/></resp></respStmt>
    <respStmt><resp>Palabras clave</resp><name/></resp></respStmt>
    <respStmt><resp>Contextualización</resp><name>Carmen Serrano Sánchez</name></resp></respStmt>
</titleStmt>
```

1.1.2. <publicationStmt>

This element contains the following parts:

- **<publisher>** contains the name of the institution responsible for the publication.
- **<address>** contains the address of the institution.
- **<pubPlace>** contains the place of publication.
- **<date>** contains the date of publication.
- **<availability>** contains information about copyright.

In most cases, the institution responsible for the publication will be CLUL. However, whenever the letters have been published before, the name of the first publisher will be referred.

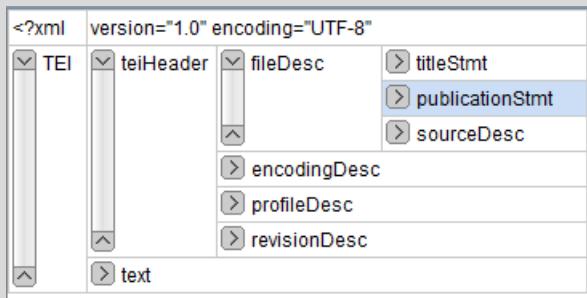
```
<publicationStmt>
    <publisher>Universidad de Extremadura</publisher>
    <pubPlace>Mérida</pubPlace>
    <date value="1999">1999</date>
    <availability status="restricted"><p>Copyright 1999, UEX</p></availability>
</publicationStmt>
```

1.1.2.1. Example

```
<publicationStmt>
    <publisher>CLUL, Centro de Linguística da Universidade de Lisboa</publisher>
    <address>
        <addrLine>Av. Professor Gama Pinto, 2</addrLine>
        <addrLine>1649-003 Lisboa-Portugal</addrLine>
    </address>
    <pubPlace>Lisboa</pubPlace>
    <date value="2013">2013</date>
    <availability status="restricted"><p>Copyright 2013, CLUL</p></availability>
</publicationStmt>
```

TEI-P5
Correspondence with the example 1.1.2.1.

Structural position of <publicationStmt>:



Encoding of the content in <publicationStmt>:

```
<publicationStmt>
    <publisher>CLUL, Centro de Linguística da Universidade de Lisboa</publisher>
    <address>
        <addrLine>Av. Professor Gama Pinto, 2</addrLine>
        <addrLine>1649-003 Lisboa-Portugal</addrLine>
    </address>
    <pubPlace>Lisboa</pubPlace>
    <date value="2013">2013</date>
    <availability status="restricted"><p>Copyright 2013, CLUL</p></availability>
</publicationStmt>
```

1.1.3. <sourceDesc>

This element contains the element <letDesc>, which includes the following parts:

- <letIdentifier> contains the archival reference of the manuscript
- <letHeading> contains the participants, place and date of the letter
- <physDesc> contains the physical description of the manuscript
- <envOcc> contains information on whether or not the letter has an envelope
- <letContents> contains information about the content of the manuscript
- <additional> contains additional information on the manuscript

1.1.3.1. <letIdentifier>

The element <letIdentifier> contains the full identification of the historical archive where the manuscript is kept. It is composed by the following parts:

- <country> refers the country where the archival institution is settled
- <settlement> refers the city where the archival institution is settled
- <institution> contains the name of the archival institution
- <repository> contains information on the repository of the archival institution
- <collection> contains information on the collection where the manuscript is filed
- <idno> contains the specific archival reference of the manuscript

- <**altName**> contains the web address to view the manuscript online (in case the archive has an online *facsimile*)

1.1.3.1.1. <**institution**>

The element <**institution**> is always followed by the attribute @**type**. The possible values for this attribute are: “nationInstitute” (national historical archives, that is, archives under the authority of a government) or “regionInstitute” (historical archives under the authority of regional or local institutions, or autonomous institutions):

```
<institution type="nationInstitute"></institution>
```

Examples:

- Archivo Histórico Nacional
- Archivo de la Real Chancillería de Valladolid
- Archivo Histórico de la Corona de Aragón
- Archivo General de Indias
- Archivo General de Simancas
- Arquivo Nacional da Torre do Tombo

```
<institution type="regionInstitute"> </institution>
```

Examples:

- Archivo del Reino de Galicia
- Archivo Histórico Provincial de Pontevedra
- Archivo Diocesano de Cuenca
- Archivo Histórico de Asturias

1.1.3.1.2. <**repository**>

The element <**repository**> is always followed by the attribute @**reg**. The value of this attribute consists of the initials of the repository name (all the initials are internal to the project):

Example 1:

```
<repository reg="CS">Casa da Suplicação</repository>
```

Example 2:

```
<repository reg="TSO">Tribunal do Santo Oficio</repository>
```

Example 3:

```
<repository reg="INQ">Inquisición</repository>
```

(para el subfondo “Inquisición” del AHN)

1.1.3.1.3. <**idno**>

The element <**idno**> gives information about the location of the letter in the archives. In this element only the invariable information within an archival reference is encoded:

Example 1:

<idno>Letra J, Mazo 136, Caja 361</idno>⁵

Example 2:

<idno>Proceso 7445</idno>

More detailed levels of organization, applicable to the same documental unit (collection etc), as well as the letter's foliation, are encoded inside the element <**locus**> (cf. 1.1.3.5.1).

This difference between <**idno**> and <**locus**> allows for the automatic search, in the TEITOK interface, upon all the letters kept in the same documental unit. All those letters will share the same information included in the element <**idno**> and will be distinguished by the information contained in the element <**locus**>.

1.1.3.1.4. <altName>

The element <**altName**> contains the web address in which the manuscript's facsimile can be found (the link makes the whole file available). The internet link is included as a value of the attribute @n. The content of the element <**xref**> is invariable:

ES
<altName><xref subcat="http://digitarq.dgarq.gov.pt/details?id=2302748">Proceso
digitalizado</xref></altName>

PT
<altName><xref subcat="http://digitarq.dgarq.gov.pt/details?id=2302748">Processo
digitalizado</xref></altName>

1.1.3.1.5. Example

Example from letter PS4001.

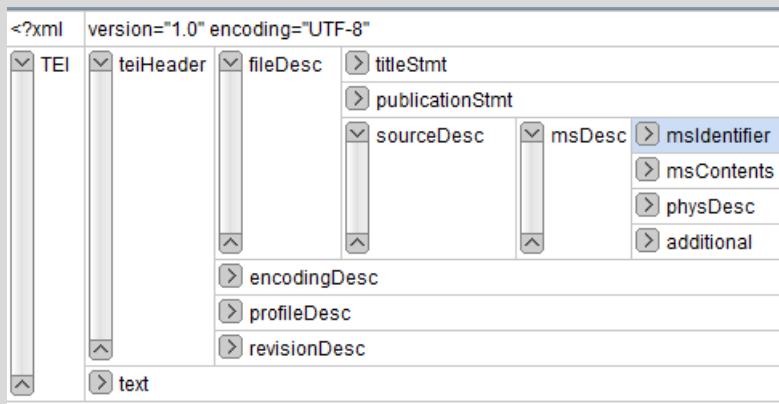
```
<letIdentifier>
  <country>España</country>
  <settlement>Madrid</settlement>
  <institution type="nation\Institute">Archivo Histórico Nacional</institution>
  <repository reg="INQ">Inquisición</repository>
  <collection>Tribunal de Distrito de la Inquisición de Toledo</collection>
  <idno>Legajo 26, Expediente 18</idno>
  <altName><xref/></altName>
</letIdentifier>
```

⁵ The content of <**idno**> is separated by commas and has an initial capital letter.

TEI-P5

Correspondence with the example in 1.1.3.1.5.

Structural position of `<msIdentifier>`:



Encoding of the content of `<msIdentifier>`:

```
<msIdentifier>
  <country>España</country>
  <settlement>Valladolid</settlement>
  <institution key="CDD.xml#NatAHN">Archivo Histórico Nacional </institution>
  <repository xml:id="INQ">Inquisición</repository>
  <collection>Tribunal de Distrito de la Inquisición de Toledo </collection>
  <idno> Legajo 26, Expediente 18</idno>
</msIdentifier>
```

1.1.3.2. `<letHeading>`

The element `<letHeading>` contains information about the letter's situational context: participants and spatial-temporal coordinates:

- `<author>` contains the name of the letter's author
- `<addressee>` contains the name of the letter's addressee
- `<placeLet>` contains the place from which the letter was sent
- `<placeDest>` contains the letter's place of destination
- `<dateLet>` contains the date of the letter

1.1.3.2.1. `<author>` and `<addressee>`

- a. The elements `<author>` and `<addressee>` have three attributes: `@attested`, `@key`, `@id`.

- b. The attribute **@attested** allows to confirm whether or not the information regarding the author is attested in the letter itself. It can have one of three possible values:

attested="yes"

The name of the author/addressee appears both in the body of the letter and in the envelope.

attested="no"

The name of the author/addressee is only known through documents appended to the proceeding, usually the court proceedings or related letters.

attested="unk"

The name of the author/addressee is unknown.

- c. The attribute **@key** links the code associated to each participant to an external database called CDD (cf. section 6). The attribute **@id** allows tracking corrections, additions and erasures -- made in the body of the letter -- to the participant's code.

```
<author attested="yes" key="CDD.xml#MH4" id="MH4">Manuel Herrera</author>
```

The attribute **@id** is only necessary inside the element **<author>**, since in the few cases in which the text is edited by the addressee, he must be treated as an annotator (cf. 1.1.3.2.1.g):

```
<addressee attested="yes" key="CDD.xml#MG12">Mateo Gutiérrez</addressee>
```

- d. If the author or the addressee are an unidentified participant, the attribute **@attested** will have the value "unk", the code of the attribute **@id** will be "AN" (for male participants) or "ANF" (for female participants), followed by the corresponding number and the name of the participant will be "Anónimo" or "Anónima", plus the corresponding number, written without spaces:

```
<author attested="unk" key="CDD.xml#AN32" id="AN32">Anónimo32</author>
<addressee attested="unk" key="CDD.xml#ANF25">Anónima25</addressee>
```

- e. If the author is anonymous, the procedure will depend on whether there is more information about him within the documental unit. If there is more information, it should be used, and the value of the attribute **@attested** will be "no"; in case there isn't, the procedure should be equal to the one applied to unidentified participants.

Example 1: (author of an anonymous letter, of whom we found information inside the proceeding)

```
<author attested="no" key="CDD.xml#ME2" id="ME2">María Estévez</author>
```

Example 2: (author of an anonymous letter, of whom we have no data)

```
<author attested="unk" key="CDD.xml#AN239" id="AN239">Anónimo239</author>
```

- f. If the letter corresponds to more than one author, there should be as many **<author>** entries, as the existent authors. Furthermore, after the second author, an attribute @n is added with the corresponding number, starting in 2:

```
<author attested="yes" key="CDD.xml#MH4" id="MH4">Manuel Herrera</author>
<author attested="yes" key="CDD.xml#FP1" id="FP1" n="2">Francisco Pérez</author>
<author attested="yes" key="CDD.xml#FV3" id="FV3" n="3">Francisca da Veiga</author>
```

The same strategy is used when there are two or more addressees:

```
<addressee attested="yes" key="CDD.xml#MH4">Manuel Herrera</addressee>
<addressee attested="yes" key="CDD.xml#FP1" n="2">Francisco Pérez</addressee>
<addressee attested="yes" key="CDD.xml#FV3" n="3">Francisca da Veiga</addressee>
```

All authors and addressees must appear in the element **<title>** and they all should have the correspondent biographical entry in the biographical file, the CDD.

- g. Optionally, the element **<respStmt>** may be used when there is a scribe or an annotator. The scribe is the person responsible for copying the letter or the person who writes the text dictated by the author; the annotator is the one who adds notes in the margins of the letter (for example, an Inquisition member or the addressee itself).

In case it is necessary to add a scribe or an annotator, that information should appear in the following way:

Example 1:

```
<respStmt><resp subcat="scribe" id="scribe"/></respStmt>
```

Example 2:

```
<respStmt><resp subcat="annotator" id="annotator"/></respStmt>
```

It is an empty element, even when the names of the scribe or the annotator are known. If they are known, they don't receive any code, but their names are mentioned in the context of the letter. If there are several scribes or annotators, they are numbered sequentially.

```
<respStmt><resp subcat="scribe" id="scribe1"/></respStmt>
<respStmt><resp subcat="scribe" id="scribe2"/></respStmt>
```

If there are annotations made by another hand in the margins of the letter, it should be mentioned in the context (cf. 1.1.3.6.3), using normalized spelling and never the actual transcription of the letter.

1.1.3.2.2. <placeLet>and <placeDest>

- a. The element <**placeLet**> gives information about the letter's place of origin. It contains the attribute @**attested**, whose values can be:

attested="yes"

The place of origin appears either in the letter or in the envelope.

attested="no"

The place of origin is unknown or we know it using information from documents appended to the letter.

- b. The element <**placeDest**> gives information about the letter's place of destination. It contains the attribute @**attested**, whose values can be:

attested="yes"

The place of destination is referred inside the letter or in the envelope.

attested="no"

The place of destination is known using information from documents appended to the letter or it is ignored.

- c. The content of elements related to geographic information, such as <**placeLet**> and <**placeDest**>, should follow the rules below. These same rules are applied to the elements <**nationality**> and <**residence**> from the CDD (cf. 6.2.10.d):

- It is filled with a reference. This reference consists on the name of the province, for Spanish data, or, for Portuguese data, on the name of the judicial district, as in the census of 1758.

Example 1:

<placeLet **attested="yes"**>Valladolid</placeLet>

Example 2:

<placeLet **attested="yes"**>Moncorvo</placeLet>

- If a more specific provenance is known, it is added using commas to separate it from the previous reference.

Example 1:

<placeLet **attested="yes"**>Valladolid, Cogeces del Monte</placeLet>

Example 2:

<placeLet **attested="yes"**>Moncorvo, Mirandela</placeLet>

- If we are dealing with a foreign country, the name of the country will be mentioned first, followed by a more specific indication (in case it is known) :

Example 1:
`<placeLet attested="yes">Francia, Paris</placeLet>`

Example 2:
`<placeLet attested="yes">Brasil, Bahia</placeLet>`

- In the case of the islands, the district is not mentioned:

Example 1:
`<placeLet attested="yes">Açores, Ilha X, local</placeLet>`

Example 2:
`<placeLet attested="yes">Madeira, local</placeLet>`

- In case of places that, at the time, belonged to a country different from their current one, there will be no reference to the country (Ceuta, Gibraltar, Goa, Damão, Diu, Malaca, Macao, ...):

Example 1:
`<nationality n="birthplace">Gibraltar</nationality>`

Example 2:
`<nationality n="birthplace">Goa</nationality>`

There is also no reference to countries that were born during the timeframe covered by *P.S. Post Scriptum* (XVI to XIX centuries): Belgium, every Popular Republics, Indonesia, ...)

Correct:

`<nationality n="birthplace">Gante</nationality>`

Incorrect:

`<nationality n="birthplace">Bélgica, Gante</nationality>`

- d. If we do not know the place of origin and/or place of destination , we write S.I. (initials of *Sem localização/ No place*):

`<placeLet attested="no">S.I.</placeLet>`

- e. The attribute **@geo** is automatically added to the elemento **<placeLet>** and its value corresponds to the geographical coordinates of that specific place, separated by a blank space:

`<placeLet attested="yes" geo="39.6156017 -2.4264300">Cuenca, Pinarejo</placeLet>`

1.1.3.2.3. <dateLet> and <date>

- a. The element <**dateLet**> gives information on the date of the letter. It contains the attribute @**attested**, whose values can be:

attested="yes"

The date appears in the letter or in the envelope.

attested="no"

Either the date is known using information from documents appended to the letter or we do not know it.

- b. The element <**dateLet**> follows the format “year.month.day”. This element contains the element <**date**>. The element <**date**> contains an attribute @**value** which corresponds to the year of the letter.

```
<dateLet attested="yes">1829.05.23<date value="1829"/></dateLet>
```

Inside the element <**dateLet**>, unknown dates are replaced by hyphens.

```
<dateLet attested="yes">1829.---<date value="1829"/></dateLet>
```

- c. The conjectured dates are encoded with square brackets, inside @**value**. Square brackets are not used inside <**dateLet**>:

```
<dateLet attested="no">1826.12.20<date value="[1826]" /></dateLet>
```

If it is not possible to conjecture a year, an interval should be conjectured⁶:

```
<dateLet attested="no">1820-1829.---<date value=" [1820-1829]" /></dateLet>
```

- d. If the letter was written along several days, the first and last days of the interval are written, separated by a hyphen:

```
<dateLet attested="yes">1824.01.18-20<date value="1824" /></dateLet>
```

- e. For the purpose of distribution by centuries, letters written in 1500 belong to the XVI century; letters written in 1600 belong to the XVII century. A letter with a conjectured date, whose interval comprehends two centuries, belongs to the first one. For example, [1690-1710] belongs to the XVII century.

1.1.3.2.4. Example

⁶ After the manuscript has been transcribed into the XML format and before the file upload to TEITOK, an attribute @when is automatically added to the <**date**> element . It contains a year value, free of square brackets (ex: @when="1824"). When a time interval is at stake, the value of the @when attribute is the middle year of the time interval.

Example from letter PS6006.

```
<letHeading>

    <author attested="yes" key="CDD.xml#GS4" id="GS4">Gregorio de Sarasa</author>

    <addressee attested="yes" key="CDD.xml#JE11" id="JE11">Juan de Ena</addressee>

    <respStmt><resp subcat="annotator" id="annotator"/></respStmt>

    <respStmt><resp subcat="scribe" id="scribe"/></respStmt>

    <placeLet attested="no" geo="41.6488226 -0.8890853">España, Zaragoza</placeLet>

    <placeDest attested="no">España, Zaragoza</placeDest>

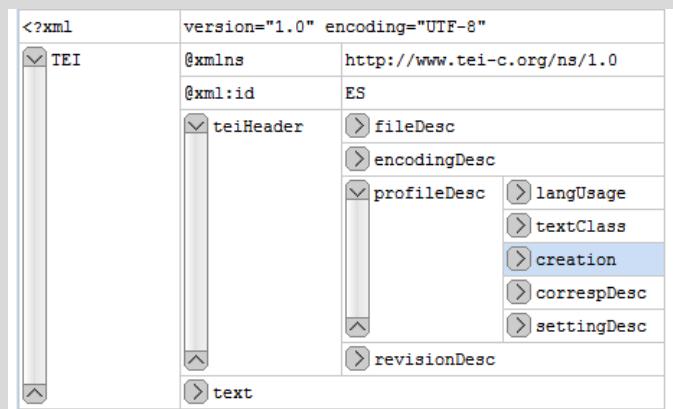
    <dateLet attested="no">1594.--.--<date value="[1594]" /></dateLet>

</letHeading>
```

TEI-P5

Correspondence with the example in 1.1.3.2.4.

Structural position of **<creation>**:



Encoding of the content of **<creation>**:

```
<creation>

    <roleName xml:id="scribe"/>

    <roleName xml:id="annotator"/>

</creation>
```

Structural position of **<correspDesc>**

```

<?xml version="1.0" encoding="UTF-8"
      xmlns="http://www.tei-c.org/ns/1.0">
<TEI>
    <@xmlns> http://www.tei-c.org/ns/1.0
    <@xml:id> ES
    <teiHeader>
        <fileDesc>
        <encodingDesc>
            <profileDesc>
                <langUsage>
                <textClass>
                <creation>
                <correspDesc> <!-- This is the node currently selected -->
                <settingDesc>
            </profileDesc>
        </encodingDesc>
    </teiHeader>
    <text>
    <correspDesc>
        <correspAction type="sent">
            <persName key="CDD.xml#GS4" xml:id="GS4" cert="high"><name>Gregorio de Sarasa</name></persName>
            <placeName confidence="0" type="origin">España, Zaragoza
            <location><geo>41.6488226 -0.8890853</geo></location>
        </placeName>
        <date confidence="0" when="1594" from="1594" to="1594" when-custom="1594"/>
        </correspAction>
        <correspAction type="received">
            <persName key="CDD.xml#JE11" xml:id="JE11" cert="high"><name>Juan de Ena</name></persName>
            <placeName confidence="0" type="destination">España, Zaragoza
            <location><geo/></location>
        </placeName>
        </correspAction>
    </correspDesc>
</text>
</TEI>

```

1.1.3.3. <physDesc>

The element <physDesc> contains information regarding the physical description of the manuscript. It contains the following elements:

- <**type**> contains information of whether it is a letter or a note
- <**support**> contains information on the physical support of the letter
- <**extent**> contains information on the dimensions of the letter.
- <**layout**> contains information about the text layout.
- <**decoration**> contains information on the existence of pictures or drawings.

- **<condition>** contains information on the state of preservation of the manuscript.

The information contained in the elements **<support>**, **<layout>** and **<condition>** is always written with initial capital letter and full stop at the end.

1.1.3.3.1. **<type>**

The element **<type>** has only two possible options: letter or note.

1.1.3.3.2. **<support>**

- a. The element **<support>** describes the support in which the letter was written. The designation in the Portuguese part of the corpus is different from the Spanish one. The table below shows a list of equivalences:

<i>Portuguese corpus</i>	<i>Spanish corpus</i>
uma folha – a full sheet of paper	un bifolio ⁷
meia folha – a folio	un folio
um quarto de folha – a quarto	un cuarto
um oitavo de folha – an octavo	un octavo
...	...

In the Spanish case, and according to the general definition, a folio is the result of folding once a standard-size sheet of paper, which remained the standard size of written documents for centuries. The recto and verso are differentiated (Cruz Mundet, 2011: 176)⁸.

- b. For all other sizes, smaller than a folio, we take as reference a sheet. A sheet folded in two is not half a folio, but a quarto, since it is the fourth part of the sheet of paper.

⁷ Altough "bifolio" and "pliego" are equivalent terms, we decided to opt for "bifolio".

⁸ Cruz Mundet, José Ramón (2011): *Diccionario de Archivística*, Madrid: Alianza Editorial.

Example 1:

ES

<support><p>un folio de papel escrito por recto y verso.</p></support>

PT

<support><p>meia folha de papel escrita no rosto e no verso.</p></support>

Example 2:

ES

<support><p>un folio de papel escrito por recto y verso y un folio de papel escrito por el recto.</p></support>

PT

<support><p>duas meias folhas de papel, a primeira escrita no rosto e no verso, a segunda, só no rosto.</p></support>

Example 3:

ES

<support><p>dos folios de papel escritos por recto y verso y un folio de papel escrito por el verso.</p></support>

PT

<support><p>duas meias folhas de papel escritas no rosto e no verso e uma terceira escrita no verso.</p></support>

Example 4:

ES

<support><p>un folio de papel doblado en cuarto, escrito por recto y verso del primer cuarto.</p></support>

PT

<support><p>meia folha de papel dobrada escrita no rosto e no verso do primeiro fólio.</p></support>

Example 5:

ES

<support><p>un cuarto de papel escrito por recto y verso.</p></support>

PT

<support><p>um quarto de folha de papel escrito no rosto e no verso.</p></support>

Example 6:

ES

<support><p>un folio de papel doblado en cuarto, escrito por recto y verso del primer cuarto y por el verso del segundo cuarto.</p></support>

PT

<support><p>meia folha de papel dobrada em quarto, escrita no rosto e verso do primeiro fólio e no verso do último.</p></support>

Example 7:

ES

<support><p>un folio de papel doblado en cuarto, escrito por todas las caras.</p></support>

PT

<support><p>meia folha de papel dobrada em quarto, escrita em todas as faces.</p></support>

Example 8:

ES

<support><p>un folio de papel doblado longitudinalmente, escrito por recto y verso.</p></support>

PT

<support><p>meia folha de papel dobrada longitudinalmente, escrita no rosto e no verso.</p></support>

Example 9:

ES

<support><p>folio doblado de manera irregular, escrito....</p></support>

PT

<support><p>meia folha de papel dobrada de maneira irregular, escrita....</p></support>

Example 10:

ES

<support><p>bifolio de papel doblado en folio, escrito por recto y verso del primer folio.</p></support>

PT

<support><p>uma folha de papel dobrada, escrita no rosto e no verso do primeiro fólio.</p></support>

Example 11:

ES

<support><p>un folio de papel doblado en cuarto, escrito por todas las caras y con sobrescrito en la última.</p></support>

PT

<support><p>meia folha de papel dobrada em quarto, escrita em todas as faces e com sobrescrito na última.</p></support>

- c. In the rare case there is a letter written in a support other than paper, the information should be as follows:

ES

<support><p>un trozo de (tipo de material) escrito por el derecho / por el revés / por los dos lados.</p></support>

PT

<support><p>um pedaço de (tipo de material) escrito na frente / no verso/ nos dois lados.</p></support>

1.1.3.3.3. **<extent>**

The element **<extent>** includes an element **<dimensions>**, which contains the dimensions of the letter. To measure the manuscript, the orientation of the writing is taken into account: the height is the line perpendicular to the text and the width is the line parallel to the orientation of the text. Both measures are expressed with three digits (in millimetres)⁹:

```
<extent>
  <dimension>
    <height units="mm">209</height><width units="mm">148</width>
  </dimensions>
</extent>
```

⁹ In the case of letters in microfilm, if there isn't a ruler on the base of the microfilm, the information about dimensions will be empty.

1.1.3.3.4. <layout>

- a. The element **<layout>**, which includes an element **<p>**, gives information on the estimated number of blank lines between the greeting form and the beginning of the text. Therefore, **<layout>** has a systematic structure as illustrated below:

ES

"sin/una/dos/tres/n línea(s) en blanco entre la fórmula de saludo y el inicio del texto".

PT:

"sem/uma-duas/três/n linha(s) em branco entre a fórmula de endereço e o início do texto".

- b. In case there is no greeting form, the reference should be the header of the letter. In this case, the systematic structure will be as follows:

ES

"sin/una/dos/tres/n línea(s) en blanco entre el encabezamiento y el inicio del texto".

PT:

"sem/uma-duas/três/n linha(s) em branco entre o cabeçalho e o início do texto".

- c. The greeting form identifies the addressee of the letter and his relationship with the sender. The header refers to any kind of information apart from the greeting form; usually, it identifies the place and/or date of the letter or contains a symbolic invocation (for instance, a cross, or the formula JMJ, standing for Jesus, Mary, Joseph).

Example 1:

Mi querido amigo,

[línea en blanco]

Te escribo estas líneas para....

<layout><p>una línea en blanco entre la fórmula de saludo y el inicio del texto.</p></layout>

Example 2:

Vigo, 15 de enero de 2013

[línea en blanco]

Te escribo estas líneas para....

<layout><p>una línea en blanco entre el encabezamiento y el inicio del texto.</p></layout>

Example 3:

Vigo, 15 de enero de 2013

[línea en blanco]

Mi querido amigo, te escribo estas líneas para....

<layout><p>sin cambio de línea entre la fórmula de saludo y el inicio del texto.</p></layout>

Example 4:

Mi querido amigo, te escribo estas líneas para...

<layout><p>sin cambio de línea entre la fórmula de saludo y el inicio del texto.</p></layout>

Example 5:

[línea en blanco]

[línea en blanco]

Te escribo estas líneas para....

<layout><p>dos líneas en blanco antes del inicio del texto.</p></layout>

Example 6:

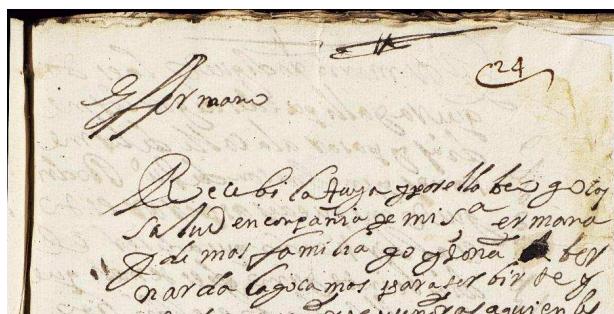
[invocación simbólica en forma de cruz]

[línea en blanco]

Te escribo estas líneas para....

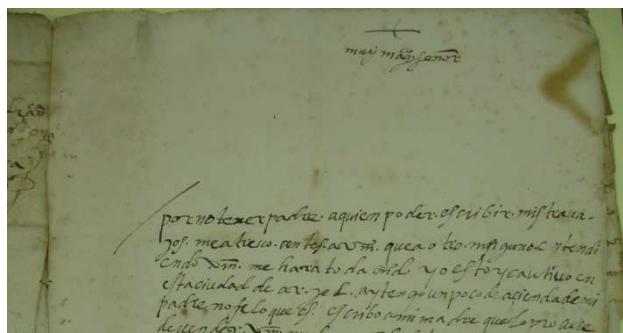
<layout><p>una línea en blanco entre la cruz del encabezamiento y el inicio del texto.</p></layout>

Illustrated examples:



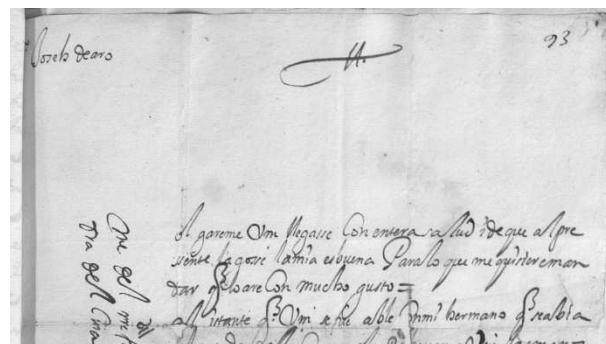
PS5013

<layout><p>dos líneas en blanco entre la fórmula de saludo y el inicio del texto.</p></layout>

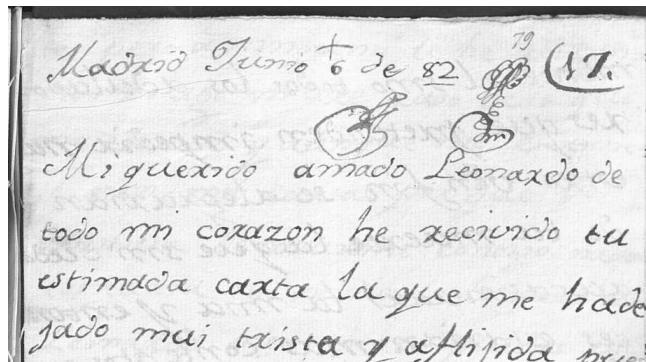


PS5012

<layout><p>ocho líneas en blanco entre la fórmula de saludo y el inicio del texto.</p></layout>

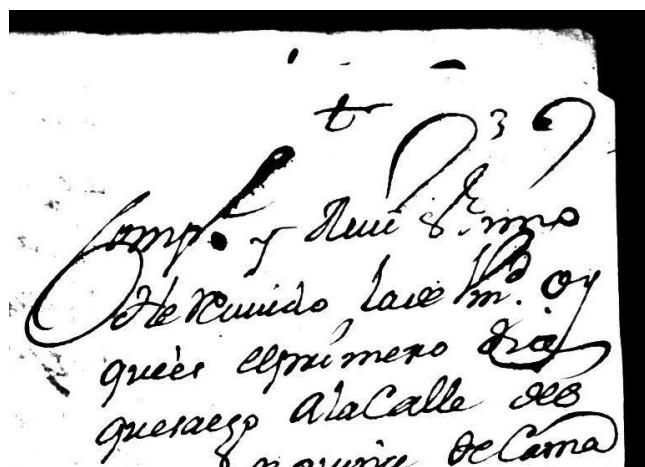


<layout><p>cinco líneas en blanco entre la cruz del encabezamiento y el inicio del texto.</p></layout>



PS5010

<layout><p>sin cambio de línea entre la fórmula de saludo y el inicio del texto.</p></layout>



PSCR5530

<layout><p>sin líneas en blanco entre la fórmula de saludo y el inicio del texto.</p></layout>

- d. If there are two greeting forms, the first one is the reference point:

Al señor Pedro Álvarez
[línea en blanco]
[línea en blanco]
Mi querido tío, te escribo estas líneas para....

<layout><p>dos líneas en blanco entre la fórmula de saludo y el inicio del texto.</p></layout>

- e. If the greeting form is written on any part of the left side margin of the folio, forming a vertical page break (a column) separated from the rest of the text (take letter PSCR1507 as an example), the procedure should be as follows:

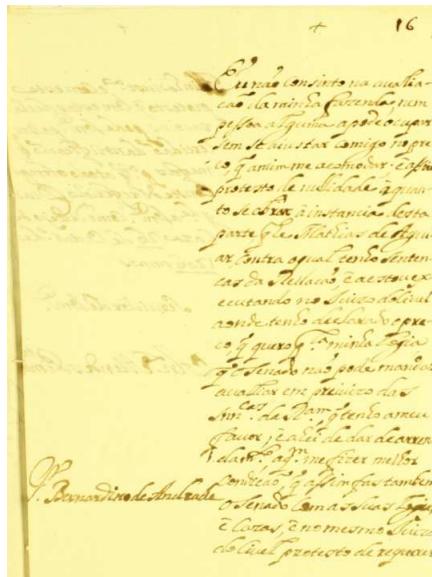
ES

<layout><p>cambio de columna entre la fórmula de saludo y el inicio del texto.</p></layout>

PT

<layout><p>mudança de coluna entre a fórmula de endereço e o início do texto.</p></layout>

Illustrated example:



PSCR1507

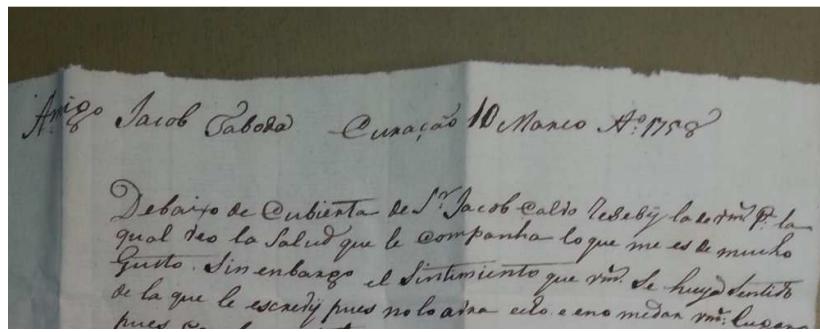
<layout><p> cambio de columna entre la fórmula de saludo y el inicio del texto.</p></layout>

- f. If there are two greeting forms separated by some blank lines, the reference should be the first greeting form and the blank lines are counted until the text begins. In this case, the formula “en blanco/em branco” (blank) is not used. In alternative, an xml comment “layout provisional/layout provisório” (provisory layout) is added, for further treatment.

Al señor Pedro Álvarez
[línea en blanco]
Vigo, 7 de octubre de 1787
[línea en blanco]
Mi querido tío, te escribo estas líneas para...

<layout><p>tres líneas entre la fórmula de saludo y el inicio del texto. </p></layout>><!--layout provisional-->

Illustrated Example:



PSCR5513

<layout><p>dos líneas entre la fórmula de saludo y el inicio del texto. </p></layout>><!--layout provisional-->

1.1.3.3.5. <decoration>

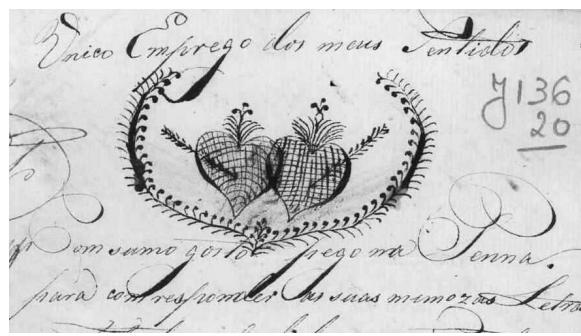
- a. The element **<decoration>** is optional and it is used to signal the presence of a drawing, a picture, or any other non alphabetic units. The element **<decoration>** always appears after **<layout>** but before **<condition>**, and has the following structure:

```

<decoration>
  <decoList>
    <decoItem id="fig1">10
      <decoDesc><p>dibujo de una rosa</p></decoDesc>
    </decoItem>
  </decoList>
</decoration>

```

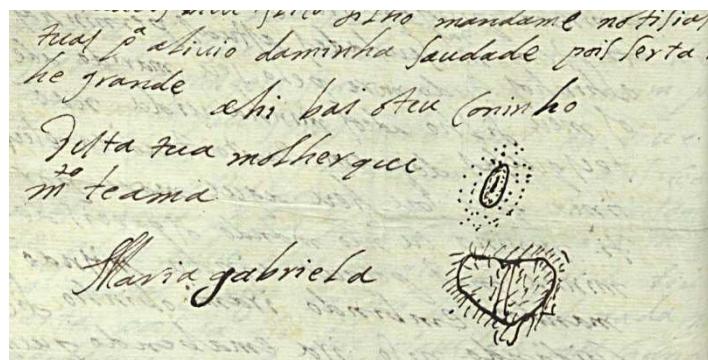
Illustrated examples:



CARDS0073

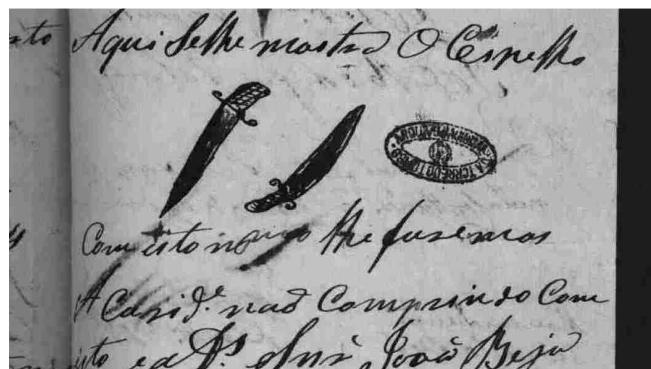
<decoDesc><p>desenho a tinta de dois corações sobrepostos, atravessados de setas e rodeados de cercadura floral.</p></decoDesc>

¹⁰ When transcribing the text, at the same place where the image appears, the following element should be added: <deco decoRef="fig1" />.



CARDS2700

<decoDesc><p>desenho de natureza erótica.</p></decoDesc>



CARDS6020

<decoDesc><p>desenho de duas espadas.</p></decoDesc>

- b. If there are several different decorations, they should be declared in different elements <decoItem>:

```

<decoration>
  <decoList>
    <decoItem id="fig1">
      <decoDesc><p>dibujo de una rosa</p></decoDesc>
    </decoItem>
    <decoItem id="fig2">
      <decoDesc><p>dibujo de un clavel</p></decoDesc>
    </decoItem>
  </decoList>
</decoration>

```

The number of declared decorations (fig 1, fig 2, etc) should be the same as the number of different types of drawing present in the text. Each type of drawing is declared only once, even if it appears several times in the same letter.

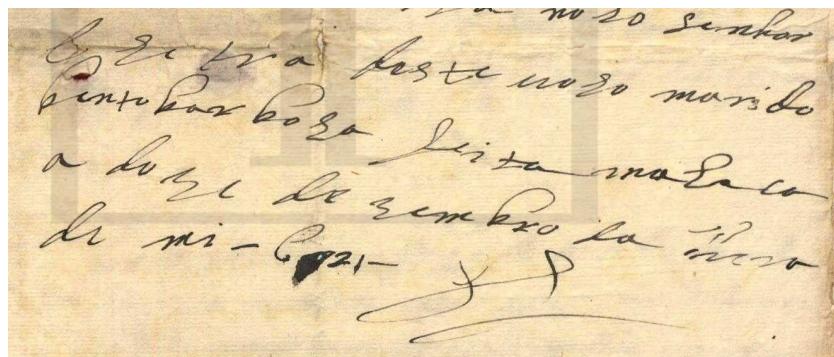
- c. If there is a symbolic invocation in the shape of a cross, the description of the drawing should contain the term “cruz” (cross):¹¹

```

<decoration>
  <decoList>
    <decoltem id="fig1">
      <decoDesc><p>cruz</p></decoDesc>
    </decoltem>
  </decoList>
</decoration>

```

- d. In some letters we find an arabesque, usually followed by a signature at the end of the letter (cf. illustration below). These cases originally represented a cross as a symbol of invocation of the Holy Trinity; then they crystallized as symbols accompanying signatures. Therefore, in these cases an image is declared in the header and is described as “symbol”:



PSCR0384

```
<decoDesc><p>símbolo</p></decoDesc>
```

- e. If it is not possible to describe the decoration, the generic term “figura” (picture) will be used.

```

<decoration>
  <decoList>
    <decoltem id="fig1">
      <decoDesc><p>figura</p></decoDesc>
    </decoltem>
  </decoList>
</decoration>

```

¹¹ The term “signo” (sign) is used for the signs appearing in the XXX of the documents; the term “invocación simbólica” (symbolic invocation) is more generic, since it can refer to the presence of a crozz or a Christogram (Chi Rho).

- f. Only the relevant drawings, figures or graphic elements are treated as decorations. Symbols working as linking marks inside the text are not treated as decorations.

1.1.3.3.6. <condition>

The element **<condition>** is optional and it is used to register any additional information on the state of preservation of the letter.

Example 1:

`<condition><p>la encuadernación esconde el inicio de las líneas del folio 15r.</p></condition>`

Example 2:

`<condition><p>la carta está bastante deteriorada a causa de la humedad.</p></condition>`

Example 3:

`<condition><p>el sobrescrito tiene restos de lacre.</p></condition>`

1.1.3.3.7. Example

Example from letter PS4001.

```

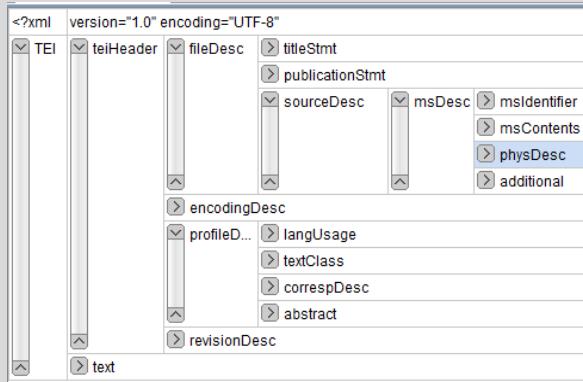
<physDesc>
  <type>carta</type>
  <support><p>un folio de papel escrito por recto y verso.</p></support>
  <extent>
    <dimensions>
      <height units="mm">304</height><width units="mm">194</width>
    </dimensions>
  </extent>
  <layout><p>tres líneas en blanco entre la fórmula de saludo y el inicio del texto.</p></layout>
  <decoration>
    <decoList>
      <decoItem id="fig1"><decoDesc><p>cruz</p></decoDesc></decoItem>
    </decoList>
  </decoration>
  <condition><p>la encuadernación esconde el final de las líneas del folio 11v.</p></condition>
</physDesc>

```

TEI-P5

Correspondence with the example in 1.1.3.3.7.

Structural position of `<physDesc>`¹²:



Encoding of the content of `<physDesc>`:

```
<physDesc>
  <objectDesc form="carta">
    <supportDesc>
      <extent>un folio de papel escrito por recto y verso.
        <dimensions scope="all" unit="mm">
          <height>304</height><width>194</width></dimensions>
      </extent>
      <condition><p>la encuadernación esconde el final de las líneas del folio 11v.</p></condition>
    </supportDesc>
    <layoutDesc>
      <layout><p> tres líneas en blanco entre la fórmula de saludo y el inicio del texto.</p></layout>
    </layoutDesc>
  </objectDesc>
  <accMat>
    <p>Sobrescrito:</p>
    <p>
      <quote xml:lang="ES"><seg type="sign">cruz</seg> <lb break="yes"/> Avisos que ave
        <lb break="no"/> ys de azer para <abbr>q</abbr> <lb break="yes"/> salga de aqui </quote>
    </p>
  </accMat>
</physDesc>
```

1.1.3.4. `<envOcc>`

The element `<envOcc>` gives information on the existence of an envelope. It contains the attribute `@occ`, which only admits one of two possible values:

```
<envOcc occ="yes"/>
  La carta tiene sobrescrito
```

```
<envOcc occ="no"/>
  La carta no tiene sobrescrito
```

¹² The element `<decoration>` is encoded within `<text>` and the content of the superscript is no longer part of `<text>`, becoming part of `<physDesc>`.

1.1.3.5. <letContents>

The element **<letContents>** has information related to the content of the text, as well as the foliation of the manuscript. It is formed by the following parts:

- **<locus>** contains the letter's folio number or numbers
- **<class>** contains the classification of the letter from different perspectives (five different categories)
- **<p></p>** contains the summary of the letter's contents

1.1.3.5.1. <locus>

The element **<locus>** includes every level of organization of an archive that concern one same document unit (collection, folio, etc...):

Example 1:

<locus>23r-v</locus>

Example 2:

<locus>Pieza 1, 23r-v</locus>

The reference to the folios of a letter takes into consideration the following indications:

- Folio written in recto and verso:

<locus>23r-v</locus>

- In case of several consecutive folios, the span is indicated:

<locus>23r-25v</locus>

- In case of non consecutive folios, the autonomous numbers are indicated:

<locus>23r, 24v y 25v</locus>

- In case of a conjecture, the folio number is written within square brackets:

<locus>[23]r-v</locus>

- In case of a conjecture involving a folio between two numbered folios, the conjectured number appears in square brackets and it is the same as the previous numbered folio, but with a letter appended to the number:

<locus>[23a]r-v</locus>

(conjectured folio between folios 23 and 24)

- f. In case the relevant folio is a loose leaf inside the proceeding, one of the previous rules is applied. Furthermore, the context of the letter should refer the fact that it is a loose leaf.

1.1.3.5.2. **<class>**

The element **<class>** is used to classify the content of the letter from different perspectives. There are four classifications:

- **<class subcat="pragmatics">** enunciation type
- **<class subcat="type">** traditional epistolary classification
- **<class subcat="linguisticSource">** linguistic keywords
- **<class subcat="socioHistoricalSource">** sociohistorical keywords

- a. The element **<class subcat="pragmatics">** contains only one from the following list of options for speech acts (the selection is mandatory):

Assertive acts		Directive acts		Compromissive acts	
ES	PT	EN	ES	PT	EN
conclusiones	conclusões	<i>conclusion</i>	aviso	aviso	<i>warning</i>
confesión	confissão	<i>confession</i>	consejo	conselho	<i>advice</i>
crítica	crítica	<i>criticism</i>	ánimo	encorajamento	<i>encouragement</i>
debate	debate	<i>debate</i>	extorsión	extorsão	<i>extortion</i>
información	informação	<i>information</i>	instrucciones	instruções	<i>instructions</i>
noticias	notícias	<i>news</i>	orden	ordem	<i>order</i>
			peticIÓN	pedido	<i>request</i>
			pregunta	pergunta	<i>question</i>
			prohibición	proibiçao	<i>prohibition</i>
			súplica	súplica	<i>supplication</i>

Expressive acts		Declarative acts	
ES	PT	EN	ES
agradecimiento	agradecimento	<i>acknowledgement</i>	acusación
elogio	elogio	<i>compliment</i>	acusaçao
presentación de	apresentação de	<i>apology</i>	declaración
disculpas	desculpas		denuncia
arrepentimiento	arrependimento	<i>regret</i>	pésame
saludos	cumprimentos	<i>greetings</i>	felicitación
despedida	despedida	<i>farewell</i>	pêsames
excusa	escusa	<i>excuse</i>	felicitação
expresión de amor	expressão de amor	<i>expression of love</i>	
expresión de odio	expressão de ódio	<i>expression of hate</i>	
expresión de pena	expressão de pena	<i>expression of sorrow</i>	
protesta	protesto	<i>protest</i>	
queja	queixa	<i>complaint</i>	
reclamación	reclamaçao	<i>reclamation</i>	
reproche	reprimenda	<i>reprimend</i>	
votos	votos	<i>wishes</i>	

- | | |
- b. The element **<class subcat="type">** contains one of four possibilities: friendship, love, family, private and anonymous (the selection is mandatory):

<class subcat="type">amistad</class>¹³

Letter between people on a friendship relationship.

<class subcat="type">amor</class>

Letter between people on a sentimental relationship.

<class subcat="type">familiar</class>

Letter between people with some level of kinship.

<class subcat="type">particular</class>

Letter between people who, not being friends, lovers or family, share some sort of complicity on everyday matters.

<class subcat="type">anónima</class>

Letter written by an author who has deliberately hidden his identity.

The fact that the participants are lovers or married to each other doesn't necessarily mean it is a love letter. The attribution of the type "love" is related to the context of the text and not to the relation between participants. Therefore, it is applied whenever the text has a sentimental content (including also negative feelings within the romantic relationship: resentment, hate, grudge, jealousy, etc.)

A letter between husband and wife whose content is not romantic is classified as "familiar"; a letter between lovers whose content is not romantic is classified as "particular".

This epistolary classification is not applied to notes.

- c. The element **<class subcat="linguisticSource">** contains one or more out of four possible options: L2, copy, hyper-segmentation, hypo-segmentation (the selection is optional):

L2:

Letter of an author who doesn't write in his native language.

(these letters are not part of the linguistic corpus)

Copia:

Letter that is not original

(these letters are not part of the linguistic corpus)

¹³ After the manuscript has been transcribed into the XML format and before the file upload to TEITOK, the "amistad/amizade" value is automatically edited and replaced by "ami.". Such replacement is justified by the need to have the same value both within Spanish and Portuguese documents (ES and PT).

Hipersegmentation

Letter in which the hand shows a clear tendency to separate parts of words.

Hiposegmentation:

Letter in which the hand shows a clear tendency to write different words together.

d. Element **<class subcat="socioHistoricalSource">**. Under construction.

e. Element **<class subcat="balancedSelection">** contains the indication on whether or not the letter is relevant for a balanced selection of documents. In such selection there is the proportion 1 letter X 1 person. The value for relevance in terms of balanced selection is "1"; the value for non-relevance is "0". This selection, which has not been applied, will be done automatically, using as criterion the number of different linguistic forms ("types") contained in the letter. This selection is carried out just before the tokenization process (cf. TEITOK below section 4.).

1.1.3.5.3. Summary of the letter

The summary of the letter is written using initial capital letter and ends with a full stop. In its composition the following aspects should be taken into account:

- Be as concise as possible.
- Start with an initial fixed structure: *El autor pide/cuenta/se queja...*¹⁴
- Write in the present tense.
- Include only information related to the letter (other type of relevant information belongs to the context).
- People who are not referred to by their name or by the expressions "author/addressee" should be referred to in a generic way.

1.1.3.5.4. Example

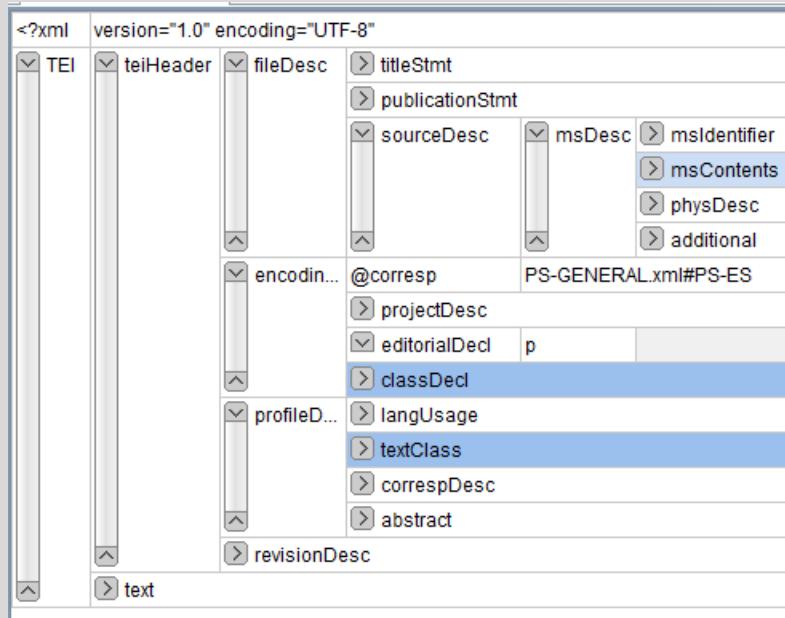
Example from letter PS4001:

```
<letContents>
<locus>11r-v</locus>
<class subcat="pragmatics">instrucciones</class>
<class subcat="type">familiar</class>
<class subcat="balancedSelection">1</class>
<class subcat="linguisticSource"/>
<class subcat="socioHistoricalSource"/>
<p>El autor indica a su hermano Diego Martínez de la Cañada qué acciones deben llevarse a cabo para que él pueda salir de la cárcel.</p>
</letContents>
```

¹⁴ The verb used should be consistent with the enunciative type of the letter. (cf. 1.1.3.5.2.a).

TEI-P5
Correspondence with the example in 1.1.3.5.4.

Structural position of **<msContents>**, **<classDecl>** and **<textClass>**:



Encoding of the content of **<msContents>**, **<classDecl>** and **<textClass>**:

```

<msContents>
    <summary>El El autor indica a su hermano Diego Martínez de la Cañada qué acciones deben llevarse
    a cabo para que él pueda salir de la cárcel.</summary>
    <msItem class="original">
        <p><locus>11r-v</locus></p>
    </msItem>
</msContents>
<physDesc/>
<additional/>
</msDesc>
</sourceDesc>
</fileDesc>
<encodingDesc>
    <projectDesc><p></p></projectDesc>
    <editorialDecl><p></p></editorialDecl>
    <classDecl>
        <taxonomy xml:id="Tip-PS">
            <bibl>Tipología PS</bibl>
            <category xml:id="pragmatics"><catDesc>instrucciones</catDesc></category>
            <category xml:id="type"><catDesc>familiar</catDesc></category>
            <category xml:id="balancedSelection"><catDesc>1</catDesc></category>
            <category xml:id="linguisticSource"><catDesc></catDesc></category>
            <category xml:id="socioHistoricalSource"><catDesc></catDesc></category>
        </taxonomy>
    </classDecl>
</encodingDesc>
<profileDesc>
    <langUsage/>
    <textClass>
        <catRef target="#pragmatics"/>
        <catRef target="#type"/>
        <catRef target="#balancedSelection"/>
        <catRef target="#linguisticSource"/>
        <catRef target="#socioHistoricalSource"/>
    </textClass>
</profileDesc>

```

1.1.3.6. <additional>

The element **<additional>** contains additional information about the letter. It is formed by the following parts:

- **<adminInfo>** contains information on the availability of the images
- **<surrogates>** contains the list of JPG files associated to each letter.
- **<note subcat="context">** contains the narration of the situational context of the letter.

1.1.3.6.1. <adminInfor>

The content of the element **<adminInfo>** is invariable, except for the name of the archival institution.

ES

```
<adminInfo>
    <availability><p>Para hacer uso de las imágenes es necesaria la autorización del
    Archivo de la Real Chancillería de Valladolid.</p></availability>
</adminInfo>
```

PT

```
<adminInfo>
    <availability><p>Reproduções não dispensam licença da DGARQ/TT.</p></availability>
</adminInfo>
```

1.1.3.6.2. <surrogates>

The element **<surrogates>** includes an element **<p>**. Inside **<p>** there is an element **<bibl>** with invariable content. The number of elements **<bibl>** corresponds to the number of image files associated to the letter:

```
<surrogates>
    <p>
        <bibl>facsímile digital guardado como fichero JPEG</bibl>
        <bibl><xref subcat="PS6143_1.JPG" /></bibl>
        <bibl><xref subcat="PS6143_2.JPG" /></bibl>
        <bibl><xref subcat="PS6143_3.JPG" /></bibl>
    </p>
</surrogates>
```

The name of the image files should follow the rules below:

- Code of the letter, followed by underscore, followed by page number (in Arabic) and file extension in capital letters. Example: PS6143_3.JPG
- If there is an envelope, number 1 is always assigned to its image file. Example: PS6143_1.JPG

- If one same image contains the envelope and a part of the text from the letter, the image should be duplicated and this information should be written inside **<support>**:

<support><p>un folio de papel doblado en cuarto, escrito por todas las caras y con sobrescrito en la última.**</p></support>**

- If one same image is shared by two different digital files, the image should be duplicated and each JPG file should have a name that matches the correspondent XML file.

1.1.3.6.3. Letter context

The element **<note subcat="context">** includes the situational context of the letter. In order to write the context, the following rules should be taken into account:

- It is written with initial capital letter and final full stop.
- Since it belongs to the textual type “narration”, verbs should be in the past tense.
- Contexts have the purpose of helping to situate the letter. Details that are not relative to the participants of the letter nor to the communicative situation should not be included.
- The transcribers’ personal judgements should be avoided.
- If parts of the proceedings transcription are included, the spelling should be normalized. Any annotation appended to the text of the letter that is included in the context should also have its spelling normalized.
- People that are mentioned in the context should be referred to by the same expression at all times (a person referred to by the name Francisco Manuel Henriques cannot be named as “o Henriques” at another moment).
- If any bibliographic reference is used to write the context, it should be listed as follows:

```

<note subcat="context"><p>Redacción del contexto...</p>
<p>Bibliografía:</p>
<p>Baião, António (1953), Episódios Dramáticos da Inquisição Portuguesa, Vol. II, Lisboa, Seara Nova (2ª ed.).</p>
<p>Rodrigues, José Francisco (1988), "Cultura e Ciência em Portugal no Século das Luzes. A Obra Matemática de José Anastácio da Cunha", Colóquio/Ciências, nº 1, Fev. 1988, pp. 74-86.</p>
</note>
```

1.1.3.6.4. Translation to English

The element **<note subcat="translations" type="English">** contains information, in English, about the content of the letter and its context. It always contains the elements **<p subcat="title">**, **<p subcat="summary">** and **<p subcat="context">**. Exceptionally, it may also contain the full translation of the letter, specially letters written by women, inside element **<p subcat="text">**:

```

<note subcat="translations" type="English">
  <p subcat="title">Love letter from Gertrudes Rosa da Conceição, prisoner, to António José Ferreira, salesman, also in jail.</p>
  <p subcat="summary">The author, who is in jail, informs her lover about the conditions she is in and about how her process is going.</p>
  <p subcat="context">Gertrudes Rosa da Conceição was accused of committing adultery with António José Ferreira, a salesman that worked for her husband, Francisco Ferreira, a grocer. She used to receive her lover in her room while her husband was away on business and this was witnessed by her servants.[...]</p>
  <p subcat="text">«I greatly esteem your health, and I hope it comes together with happiness, as I wish for you.</p><p>I should have written earlier, but I haven't been able to, because, after I got up, I got a cold.
  [...] When you wish to write, if possible, you can give it to João from our room because he is very competent. He doesn't send them unless they are sewn to a package, because it's safer than sending the boy here, since he is very trickster, he says unthinkable things. I'm quite afraid of him. So, goodbye, see you. From the one who is and will be.»</p>
</note>
```

1.1.3.6.5. Example

Example from CARDS0006:

```
<additional>
  <adminInfo>
    <availability><p>Reproduções não dispensam licença da DGARQ/TT.</p></availability>
  </adminInfo>
  <surrogates>
    <p>
      <bibl>facsimile digital guardado como ficheiro JPEG.</bibl>
      <bibl><xref subcat="CARDS0006_1.JPG"/></bibl>
      <bibl><xref subcat="CARDS0006_2.JPG"/></bibl>
      <bibl><xref subcat="CARDS0006_3.JPG"/></bibl>
    </p>
  </surrogates>
  <note subcat="context"><p>Francisco Ferreira, merceeiro, acusa a mulher de adultério. Quando o marido estava fora (testemunharam os criados), Gertrudes acolhia no seu quarto António José Ferreira, caixeiro do marido. Como sinal de presença, António José Ferreira atirava moedas à janela, e então Gertrudes abria-lhe a porta que dava entrada para o escritório. As cartas trocadas entre os amantes eram levadas por um moço de recados que foi espancado por António José Ferreira assim que se ouviram os primeiros rumores do crime. No dia de Entrudo, Gertrudes mandou toda a família e os criados para casa de um vizinho para poder receber o seu amante. Porém, uma das criadas precisou de voltar a casa e encontrou-os em «ações indecentes». Mesmo depois de presos, Gertrudes e António José continuaram a trocar correspondência, agora na cadeia do Limoeiro. As cartas foram apreendidas e nelas se insinuavam projetos de atentar contra a vida de Francisco Ferreira, o que, de facto, aconteceu: o irmão da ré, identificado como «marujo», atacou o cunhado com «huma estocada, ou punhalada, deixando-o por morto». No entanto, Francisco Ferreira escapou do que durante muito tempo se considerou ser uma «ferida mortal». Consta do processo que a ré foi absolvida por não haver provas suficientes e se supor que tudo tivesse sido urdido pelo marido.</p></note>
  <note subcat="translations" type="English">
    <p subcat="title">Love letter from Gertrudes Rosa da Conceição, prisoner, to António José Ferreira, salesman, also in jail.</p>
    <p subcat="summary">The author, who is in jail, informs her lover about the conditions she is in and about how her process is going.</p>
    <p subcat="context">Gertrudes Rosa da Conceição was accused of committing adultery with António José Ferreira, a salesman that worked for her husband, Francisco Ferreira, a grocer. She used to receive her lover in her room while her husband was away on business and this was witnessed by her servants. Even after the lovers were both in prison, they continued to exchange letters, in which there were signs of an arrangement to kill the husband. In fact, Gertrudes' brother, known as the «sailor», did hurt his brother-in-law with a dagger, but, in spite of the serious wounds, Francisco Ferreira survived. Gertrudes ended up acquitted from this accusation, for there was not enough evidence against her.</p>
    <p subcat="text">«I greatly esteem your health, and I hope it comes together with happiness, as I wish for you.</p><p>I should have written earlier, but I haven't been able to, because, after I got up, I got a cold. [...] When you wish to write, if possible, you can give it to João from our room because he is very competent. He doesn't send them unless they are sewn to a package, because it's safer than sending the boy here, since he is very trickster, he says unthinkable things. I'm quite afraid of him. So, goodbye, see you. From the one who is and will be.»</p></note>
  </additional>
```

1.2. <encodingDesc>

This element contains an attribute @key which points to an external document. That document contains two notes, one in Spanish and another in Portuguese, with the description of the Project Post Scriptum and its editorial procedures. The attribute @key contains a value that corresponds to the Spanish or the Portuguese document, accordingly. The rest of the information contained in <encodingDesc> is invariable:

ES
<encodingDesc key="PS-GENERAL.xml#PS-ES">

PT
<encodingDesc key="PS-GENERAL.xml#PS-PT">

1.3. <profileDesc>

This element contains an element <langUsage> to describe the language used in the manuscript, which will be Spanish or Portuguese:

ES

```
<profileDesc>
    <langUsage><language subcat="ES"/></langUsage>
</profileDesc>
```

PT

```
<profileDesc>
    <langUsage><language subcat="PT"/></langUsage>
</profileDesc>
```

1.4. <revisionDesc>

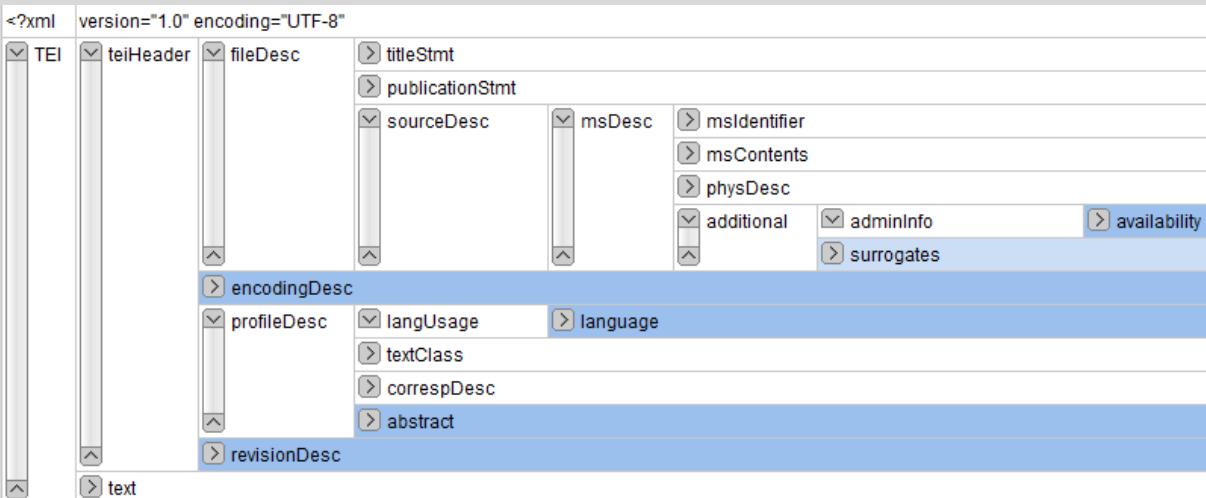
This element contains an element <change> which works as log of the different modifications and revisions done to the XML file. It is composed by an attribute @when that includes the edition date (with the format year-month-day). This element is filled with the name of the editor and the nature of the editing task:

```
<change when="2014-06-26">Gael Vaamonde: Codificación XML y contextualización</change>
<change when="2015-02-17">Rita Marquillas: Revisión de la transcripción.</change>
```

TEI-P5

Correspondence with the examples in 1.1.2.6.5., 1.2., 1.3. y 1.4.

Structural position of <availability>, <encodingDesc>, <language>, <surrogates>, <settingDesc> and <revisionDesc>¹⁵:



Encoding of the content of <availability>, <encodingDesc>, <language>, <surrogates>, <settingDesc> and <revisionDesc>:

```

<additional>
  <adminInfo>
    <availability><p>Reproduções não dispensam licença da DGARQ/TT.</p></availability>
  </adminInfo>
  <surrogates>
    <ref/>
    <cit><bibl>facsímile digital guardado como ficheiro JPEG</bibl></cit>
  </surrogates>
  </additional>
</msDesc>
</sourceDesc>
</fileDesc>
<encodingDesc corresp="PS-GENERAL.xml#PS-PT">
  <projectDesc><p/></projectDesc>
  <editorialDecl><p/></editorialDecl>
  <classDecl/>
</encodingDesc>
<profileDesc>
  <langUsage><language ident="PT"/></langUsage>
  <textClass/>
  <langUsage/>
  <textClass/>
  <correspDesc/>
  <settingDesc><p>Francisco Ferreira, merceiro, acusa a mulher de adultério. Quando o marido estava fora (testemunharam os criados), Gertrudes acholia no seu quarto António José Ferreira, caixeiro do marido. Como sinal de presença, António José Ferreira atirava moedas à janela, e então Gertrudes abria-lhe a porta que dava entrada para o escritório. As cartas trocadas entre os amantes eram levadas por um moço de recados que foi espancado por António José Ferreira assim que se ouviram os primeiros rumores do crime. No dia de Entrudo, Gertrudes mandou toda a família e os criados para casa de um vizinho para poder receber o seu amante. Porém, uma das criadas precisou de voltar a casa e encontrou-os em «ações indecentes». Mesmo depois de presos, Gertrudes e António José continuaram a trocar correspondência, agora na cadeia do Limoeiro. As cartas foram apreendidas e nelas se insinuavam projetos de atentar contra a vida de Francisco Ferreira, o que, de facto, aconteceu: o irmão da ré, identificado como «marujo», atacou o cunhado com «huma estocada, ou punhalada, deixando-o por morto». No entanto, Francisco Ferreira escapou do que durante muito tempo se considerou ser uma «ferida mortal». Consta do processo que a ré foi absolvida por não haver provas suficientes e se supor que tudo tivesse sido urdido pelo marido.</p>
  </settingDesc>
</profileDesc>
<revisionDesc>
  <change when="2007-12">Cristina Albino: Codificação XML</change>
  <change when="2012-11">Ana Luísa Costa: Revisão da leitura</change>

```

¹⁵ The reference to JPG files becomes part of the element <text>.

```

<change when="2013-01">Clara Pinto: Tradução</change>
<change when="2013-10">Rita Marquilhas: Importação para coleção PS, Post Scriptum, ERC, European Research Council - 7FP/ERC Advanced Grant 2011 GA295562</change>
<change when="2015-09-29">Fernanda Pratas: Revisão e sumário em inglês</change>
<change xml:lang="EN">
<p sameAs="title">love letter from Gertrudes Rosa da Conceição, prisoner, to António José Ferreira, salesman, also in jail.</p>
<p sameAs="summary">The author, who is in jail, informs her lover about the conditions she is in and about how her process is going.</p>
<p sameAs="settingDesc">Gertrudes Rosa da Conceição was accused of committing adultery with António José Ferreira, a salesman that worked for her husband, Francisco Ferreira, a grocer. She used to receive her lover in her room while her husband was away on business and this was witnessed by her servants. Even after the lovers were both in prison, they continued to exchange letters, in which there were signs of an arrangement to kill the husband. In fact, Gertrudes' brother, known as the «sailor», did hurt his brother-in-law with a dagger, but, in spite of the serious wounds, Francisco Ferreira survived. Gertrudes ended up acquitted from this accusation, for there was not enough evidence against her.</p>
<p sameAs="text">«I greatly esteem your health, and I hope it comes together with happiness, as I wish for you.</p><p>I should have written earlier, but I haven't been able to, because, after I got up, I got a cold. [...] When you wish to write, if possible, you can give it to João from our room because he is very competent. He doesn't send them unless they are sewn to a package, because it's safer than sending the boy here, since he is very trickster, he says unthinkable things. I'm quite afraid of him. So, goodbye, see you. From the one who is and will be.»</p></change>
</revisionDesc>

```

1.5. Full example of <teiHeader>

Example from letter PS4127.

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="CardsFly1.xsl"?>
<!DOCTYPE TEI.2 PUBLIC "-//CTB//DTD Dalf 1.0 (based on TEI)//NL" "dalf-tt.dtd">
<TEI.2>
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>1650. Carta de Juan Carpio Lozano, zapatero, para Juan de Palomares, zapatero.</title>
      <sponsor>CLUL, Centro de Linguística da Universidade de Lisboa</sponsor>
      <funder>ERC, European Research Council - 7FP/ERC Advanced Grant 2011 GA295562</funder>
      <respStmt><resp subcat="project"><name id="PS">PS, Post Scriptum</name></resp></respStmt>
      <respStmt><resp subcat="coordination"><name id="RM">Rita Marquilhas</name></resp></respStmt>
      <respStmt><resp subcat="transcription"><name id="CSS">Carmen Serrano Sánchez</name></resp></respStmt>
      <respStmt><resp subcat="posannotation"><name/></resp></respStmt>
      <respStmt><resp subcat="parsing"><name/></resp></respStmt>
      <respStmt><resp subcat="modernizedEdition"><name id="GV">Gael Vaamonde</name></resp></respStmt>
      <respStmt><resp subcat="revision"><name>Gael Vaamonde</name></resp></respStmt>
      <respStmt><resp subcat="translation"><name/></resp></respStmt>
      <respStmt><resp subcat="keywords"><name/></resp></respStmt>
      <respStmt><resp subcat="contextualization"><name>Carmen Serrano Sánchez</name></resp></respStmt>
    </titleStmt>
    <publicationStmt>
      <publisher>CLUL, Centro de Linguística da Universidade de Lisboa</publisher>
      <address>
        <addrLine>Av. Professor Gama Pinto, 2</addrLine>
        <addrLine>1649-003 Lisboa-Portugal</addrLine>
      </address>
      <pubPlace>Lisboa</pubPlace>
      <date value="2014">2014</date>
      <availability status="restricted"><p>Copyright 2014, CLUL</p></availability>
    </publicationStmt>
    <sourceDesc>
      <letDesc>
        <letIdentifier>
          <country>España</country>
          <settlement>Madrid</settlement>
          <institution type="nationInstitute">Archivo Histórico Nacional</institution>
          <repository reg="INQ">Inquisición</repository>
          <collection>Tribunal de Distrito de la Inquisición de Toledo</collection>
          <idno>Legajo 23, Expediente 14</idno>
          <altName><xref/></altName>
        </letIdentifier>
        <letHeading>
          <author attested="yes" key="CDD.xml#JCL3" id="JCL3">Juan Carpio Lozano</author>
          <addressee attested="yes" key="CDD.xml#JP12" id="JP12">Juan de Palomares</addressee>
          <placeLet attested="no">Toledo</placeLet>
          <placeDest attested="no">Madrid</placeDest>
          <dateLet attested="yes">1650.08.23<date value="1650"/></dateLet>
        </letHeading>
        <physDesc>
          <type>carta</type>
          <support><p>un bifolio de papel doblado en folio, escrito por recto y verso del primer folio y por el recto del segundo folio.</p></support>
          <extent>
            <dimensions><height units="mm">315</height><width units="mm">213</width></dimensions>
          </extent>
          <layout><p>seis líneas en blanco entre la fórmula de saludo y el inicio del texto.</p></layout>
        </physDesc>
      </letDesc>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

```

<decoration>
  <decoList>
    <decoItem id="fig1"><decoDesc><p>cruz</p></decoDesc></decoItem>
  </decoList>
</decoration>
<condition><p></p></condition>
</physDesc>
<envOcc occ="no"/>
<letContents>
  <locus>10r-11r</locus>
  <class subcat="pragmatics">instrucciones</class>
  <class subcat="type">amistad</class>
  <class subcat="balancedSelection">1</class>
  <class subcat="linguisticSource"/>
  <class subcat="socioHistoricalSource"/>
  <p>El autor da instrucciones a su amigo Juan de Palomares para que acompañe a María Rodríguez a  
retractarse de su declaración, en la que afirmaba que estaba casada con él.</p>
</letContents>
<additional>
  <adminInfo>
    <availability><p>Para hacer uso de las imágenes es necesaria la autorización del Archivo Histórico  
Nacional.</p></availability>
  </adminInfo>
  <surrogates>
    <p>
      <bibl>facsímile digital guardado como fichero JPEG</bibl>
      <bibl><xref subcat="PS4127_1.JPG"/></bibl>
      <bibl><xref subcat="PS4127_2.JPG"/></bibl>
      <bibl><xref subcat="PS4127_3.JPG"/></bibl>
    </p>
  </surrogates>
  <note subcat="context"><p>El reo de este proceso era Juan Carpio Lozano, zapatero. Fue acusado del  
delito de bigamia por haberse casado dos veces: primero con Ana Gómez en Toledo y, después, con María  
Rodríguez, viuda de Francisco Ramiro Luengo, en Madrid en 1647. Las cartas que escribió el reo a su  
amigo Juan de Palomares y a su primera mujer fueron incorporadas al proceso como prueba. Pero,  
finalmente, la causa fue suspendida, al retractarse Ana Gómez de su declaración afirmando que, en  
realidad, no estaba casada con Juan Carpio Lozano, sino amancebada.</p></note>
  <note subcat="translations" type="English"><p subcat="title"/><p subcat="summary"/><p  
subcat="context"/></note>
</additional>
<letDesc>
</sourceDesc>
</fileDesc>
<encodingDesc key="PS-GENERAL.xml#PS-ES">
  <projectDesc><p></p></projectDesc>
  <editorialDecl><p></p></editorialDecl>
</encodingDesc>
<profileDesc>
  <langUsage><language subcat="ES"/></langUsage>
</profileDesc>
<revisionDesc>
  <change when="2014-05">Carmen Serrano Sánchez: Codificación XML y contextualización</change>
  <change when="2015-03">Gael Vaamonde: Edición modernizada</change>
</revisionDesc>
</teiHeader>
<text id="PS4127">
  <body>
    <p></p>
  </body>
</text>
</TEI.2>

```

TEI-P5
Correspondence with the example in 1.5.

```

<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:id="ES">
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>1650. Carta de Juan Carpio Lozano, zapatero, para Juan de Palomares, zapatero.</title>
      <sponsor>CLUL, Centro de Linguística da Universidade de Lisboa</sponsor>
      <funder>ERC, European Research Council - 7FP/ERC Advanced Grant 2011 GA295562</funder>
      <respStmt><resp><name xml:id="PS">PS, Post Scriptum</name></resp></respStmt>
      <respStmt><resp>Coordinación</resp><name xml:id="RM">Rita Marquilhas</name></resp></respStmt>
      <respStmt><resp>Transcripción</resp><name xml:id="CSS">Carmen Serrano Sánchez</name></resp></respStmt>
      <respStmt><resp>Anotación POS</resp><name/></resp></respStmt>
      <respStmt><resp>Parsing</resp><name/></resp></respStmt>
      <respStmt><resp>Edición modernizada</resp><name>Carmen Serrano Sánchez</name></resp></respStmt>
      <respStmt><resp>Revisión</resp><name>Gael Vaamonde</name></resp></respStmt>
      <respStmt><resp>translations</resp><name/></resp></respStmt>
      <respStmt><resp>Palabras-clave</resp><name/></resp></respStmt>
      <respStmt><resp>Contextualización</resp><name>Carmen Serrano Sánchez</name></resp></respStmt>
    </titleStmt>
    <publicationStmt>
      <publisher>CLUL, Centro de Linguística da Universidade de Lisboa</publisher>
      <address>
        <addrLine>Av. Professor Gama Pinto, 2</addrLine>
        <addrLine>1649-003 Lisboa-Portugal</addrLine>
      </address>
      <pubPlace>Lisboa</pubPlace>
      <date when="2014">2014</date>
      <availability status="restricted"><p>Copyright 2014, CLUL</p></availability>
    </publicationStmt>
    <sourceDesc>
      <msDesc>
        <msIdentifier>
          <country>España</country>
          <settlement>Madrid</settlement>
          <institution key="CDD.xml#NatAHN">Archivo Histórico Nacional</institution>
          <repository xml:id="INQ">Inquisición</repository>
          <collection>Tribunal de Distrito de la Inquisición de Toledo</collection>
          <idno>Legajo 23, Expediente 14</idno>
        </msIdentifier>
        <msContents>
          <summary>El autor da instrucciones a su amigo Juan de Palomares para que acompañe a María Rodríguez a retractarse de su declaración, en la que afirmaba que estaba casada con él.</summary>
          <msItem class="original">
            <p><locus>10r-11r</locus></p>
          </msItem>
        </msContents>
        <physDesc>
          <objectDesc form="carta">
            <supportDesc>
              <extent>
                <dimensions scope="all" unit="mm">
                  <height>315</height>
                  <width>213</width>
                </dimensions>
              </extent>
              <condition></condition>
            </supportDesc>
            <layoutDesc>
              <layout><p>seis líneas en blanco entre la fórmula de saludo y el inicio del texto.</p></layout>
            </layoutDesc>
          </objectDesc>
        <accMat>
          <p>Sobrescrito:</p>
          <p><quote xml:lang="ES"/></p>
          <p>Matasellos:</p>
          <p><quote xml:lang="ES"/></p>
        </accMat>
      </msDesc>
    </sourceDesc>
    <additional>
      <adminInfo>
        <availability><p>Para hacer uso de las imágenes es necesaria la autorización del Archivo Histórico</p></availability>
      </adminInfo>
    </additional>
  </fileDesc>
</teiHeader>
<text>
  <body>
    <div>
      <div>
        <div>
          <div>
            <div>
              <div>
                <div>
                  <div>
                    <div>
                      <div>
                        <div>
                          <div>
                            <div>
                              <div>
                                <div>
                                  <div>
                                    <div>
                                      <div>
                                        <div>
                                          <div>
                                            <div>
                                              <div>
                                                <div>
                                                  <div>
                                                    <div>
                                                      <div>
                                                        <div>
                                                          <div>
                                                            <div>
                                                              <div>
                                                                <div>
                                                                  <div>
                                                                    <div>
                                                                      <div>
                                                                        <div>
                                                                          <div>
                                                                            <div>
                                                                              <div>
                                                                                <div>
                                                                                  <div>
                                                                                    <div>
                                                                                      <div>
                                                                                        <div>
              </div>
            </div>
          </div>
        </div>
      </div>
    </div>
  </body>
</text>

```

```

Nacional.</p></availability>
</adminInfo>
<surrogates>
<ref/>
<cit><bibl>facsímile digital guardado como fichero JPEG</bibl></cit>
</surrogates>
</additional>
</msDesc>
</sourceDesc>
</fileDesc>
<encodingDesc corresp="PS-GENERAL.xml#PS-ES">
<projectDesc><p></p></projectDesc>
<editorialDecl><p></p></editorialDecl>
<classDecl>
<taxonomy xml:id="Tip-PS">
<bibl>Tipología PS</bibl>
<category xml:id="pragmatics"><catDesc>instrucciones</catDesc></category>
<category xml:id="type"><catDesc>amistad</catDesc></category>
<category xml:id="linguisticSource"><catDesc/></category>
<category xml:id="socioHistoricalSource"><catDesc/></category>
<category xml:id="balancedSelection"><catDesc>1</catDesc></category>
</taxonomy>
</classDecl>
</encodingDesc>
<profileDesc>
<langUsage><language ident="ES"/></langUsage>
<textClass>
<catRef target="#pragmatics"/>
<catRef target="#type"/>
<catRef target="#linguisticSource"/>
<catRef target="#socioHistoricalSource"/>
<catRef target="#balancedSelection"/>
</textClass>
<creation>
<persName type="author" key="CDD.xml#JCL3" xml:id="JCL3" cert="high">
<name>Juan Carpio Lozano</name>
</persName>
<persName type="addressee" key="CDD.xml#JP12" cert="high">
<name>Juan de Palomares</name>
</persName>
<placeName confidence="0" type="origin">Toledo</placeName>
<placeName confidence="0" type="destination">Madrid</placeName>
<date confidence="1" when="1650-08-23" from="1650-08-23" to="1650-08-23" when-custom="1650"/>
</creation>
<settingDesc><p>El reo de este proceso era Juan Carpio Lozano, zapatero. Fue acusado del delito de bigamia por haberse casado dos veces: primero con Ana Gómez en Toledo y, después, con María Rodríguez, viuda de Francisco Ramiro Luengo, en Madrid en 1647. Las cartas que escribió el reo a su amigo Juan de Palomares y a su primera mujer fueron incorporadas al proceso como prueba. Pero, finalmente, la causa fue suspendida, al retractarse Ana Gómez de su declaración afirmando que, en realidad, no estaba casada con Juan Carpio Lozano, sino amancebada.</p>
</settingDesc>
</profileDesc>
<revisionDesc>
<change when="2014-05">Carmen Serrano Sánchez: Codificación XML y contextualización</change>
<change when="2015-03">Carmen Serrano Sánchez: Modernización</change>
<change xml:lang="EN">
<p sameAs="title"></p>
<p sameAs="summary"></p>
<p sameAs="settingDesc"></p>
</change>
</revisionDesc>
</teiHeader>
<text ml:id="PS4127">
<body>
<p></p>
</body></text>
</TEI>

```

2. The Text: <text>

The element <text> contains the textual data of the manuscript. It is composed by two elements.

- <envelope> contains the text of the envelope.
- <body> contains the text of the letter.

The element <text> contains an attribute @id which contains the code of the letter:

```
<text id="PS6234"/>
```

2.1. <envelope>

- This element contains the transcription of the text included in the envelope of the letter. It is accompanied by the attribute @type, whose value is always "superscript". The structure of the element <envelope> is the following:

```
<envelope type="superscript">
  <envPart>
    <address type="receiver">
      <addrLine></addrLine>
      <addrLine></addrLine>
      <addrLine></addrLine>
    </address>
  </envPart>
</envelope>
```

- The element <address> is followed by the attribute @type. This attribute can have one of the following two values, usually the first one:

```
<address type="receiver">
  Fot the addressee identification.
```

```
<address type="sender">
  Fort he sender identification.
```

- The line breaks inside the envelope are encoded with <addrLine>.
- Inside the element <envelope> abbreviations are not expanded, but simply marked with the element <abbr>. The tag to mark the beginning of vertical text, <lb line_rend="verticalText"/>, is not used here either.

2.2. <body>

- This element contains the transcription of the text of the letter. It is tipically composed by opener, one or more paragraphs, closer and postscript:

```

<body>
  <opener></opener>
  <p></p>
  <closer></closer>
  <ps></ps>
</body>

```

- b. The element **<body>** is followed by the element **<pb/>**, which indicates the beginning of the folio. The element **<pb/>** contains the attribute **@n**, whose value corresponds to the number of the folio that is going to be transcribed:

```
<body><pb n="16r"/>
```

2.2.1. **<opener>** and **<closer>**

- a. The textual parts whose function is to open and close the text of the letter are encoded with the elements **<opener>** and **<closer>**, accordingly. These parts are only signalled when graphically separated from the body of the text. They are, therefore, both textual and graphic elements.

The **<opener>** is typically formed by the following elements, separated by punctuation and **<lb/>** whenever the manuscript's layout justifies it:

```

<opener>
  <letAddress>Lisboa</letAddress>
  <letDate>7 de febrero</letDate> <lb/>
  <salute>Muy querido amigo y señor</salute>
</opener>

```

The **<closer>** is typically formed by the following elements, separated by punctuation and and **<lb/>** whenever the manuscript's layout justifies it:

```

<closer>
  <letAddress>Lisboa</letAddress>
  <letDate>7 de febrero de 1766</letDate> <lb/>
  <salute>Tu querido hermano</salute> <lb/>
  <signed>Pedro Álvarez</signed>
</closer>

```

- b. The elements **<letAddress>**, **<letDate>**, **<salute>** and **<signed>** include the letter's formulaic text and are defined according to their content. They are, therefore, semantic elements.
- c. The formulaic parts which are not graphically separated from the body of the text are not part of the **<opener>** nor of the **<closer>**. They are transcribed inside **<p></p>**:

Example 1:

Lisboa, 7 de febrero

[línea en blanco]

[línea en blanco]

Muy querido amigo y señor. Te escribo esta carta...

```
<opener>
  <letAddress>Lisboa</letAddress>,
  <letDate>7 de febrero</letDate>
</opener>
<p><salute>Muy querido amigo y señor</salute>. Te escribo esta carta...</p>
```

Example 2:

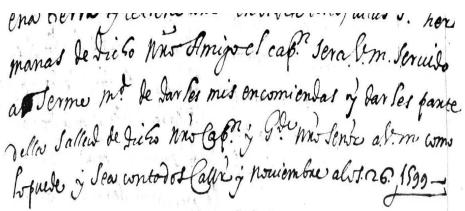
... Dios te guarde muchos años. Adiós. Lisboa, 7 de febrero de 1765

[línea en blanco]

Pedro Álvarez

```
<p>... Dios te guarde muchos años. Adiós. <letAddress>Lisboa,</letAddress> <letDate>7 de
febrero de 1765</letDate></p>
<closer>
  <signed>Pedro Álvarez</signed>
</closer>
```

Illustrated examples:

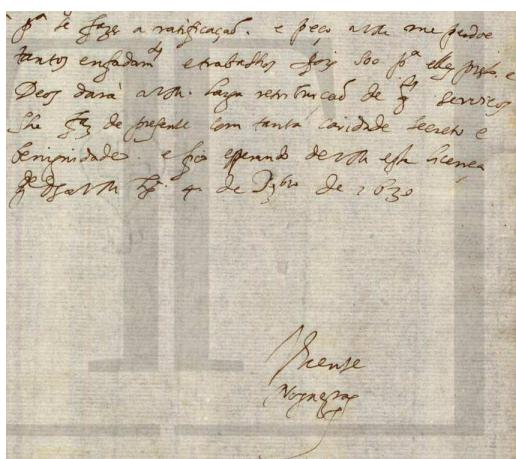


en la tierra y le tiene arta enbidia dello, a llas stas her
manas de dicho nro Amigo el capn sera Vm servido
aseme md de darles mis encomiendas y darles parte
dello sallud de dicho nro capn y Gde nro señor a Vm como
lo puede y sea con todos Caller y noviembre a los 26 de 1599

ena tierra y le tiene arta enbidia dello, a llas stas her
manas de dicho nro Amigo el capn sera Vm servido
aseme md de darles mis encomiendas y darles parte
dello sallud de dicho nro capn y Gde nro señor a Vm como
lo puede y sea con todos Caller y noviembre a los 26 de 1599

PSCR6143

Example 1: <letAddress> and <letDate> inside <p>. There is no <closer>



Deos dará a VM larga retribuição de qtos serviços
lhe faz de presente com tanta caridade secreto e
benignidade. e fico esperando de VM esta licença
gde ds a VM Lxa 4 de Dezbro 1630

Deos dará a VM larga retribuição de qtos serviços
lhe faz de presente com tanta caridade secreto e
benignidade. e fico esperando de VM esta licença
gde ds a VM Lxa 4 de Dezbro 1630

Vicente
Nogueira

PSCR1384

Example 2: <letAddress> and <letDate> inside <p>, and <signed> inside <closer>

el cap...
 son soberanos del Cnra y v...
 que la constitucion es cosa
 del infierno y condenados
 los que la queremos. Dios
 quiera que VS lo componga
 luego y le guarde ms as.
 Badocondes a 25 de Nobienbre
 de 1822

que la constitucion es cosa
 del infierno y condenados
 los que la queremos. Dios
 quiera que VS lo componga
 luego y le guarde ms as.

Badocondes a 25 de Nobienbre
de 1822.

PSCR6882

Example 3: <letAddress> and <letDate> inside <closer>

Deseo ocasion de emplearme
 en su obsequio, a quien me
 ofrezco con la mayor aten-
 sion, y respeto. Barna
 y Marzo 22. de 86.
 J. S. M. B. Fernando
 La Moga Pbro

deseo ocasion de emplearme
 en su obsequio, a quien me
 ofrezco con la mayor aten-
 sion y respeto Barna
 y Marzo 22. de 86.

Q S M B Fernando
 de la Moga Pbro

PSCR7508

Example 4: <letAddress> and <letDate> inside <p>, and <salute> and <signed> inside <closer>

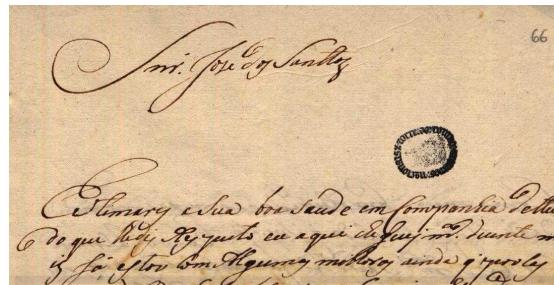
Agosto 19 de 1712 — 14
 Señor Juan Perez mi amo me man-
 da diga A Vmd le aga favor de
 pasar con el mozo a la Botica

Agosto 19 de 1712

Señor Juan Perez mi amo me man-
 da diga A Vmd le aga favor de
 pasar con el mozo a la Botica

PSCR6658

Example 5: <letDate> inside <opener>, and <salute> inside <p>

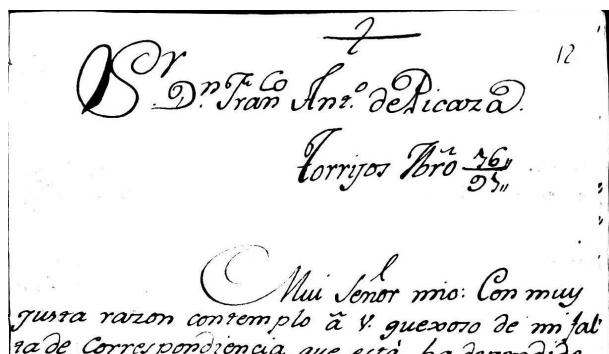


Snr Jose dos Santos

Estimarei a sua boa saude em companhia de ttu
do que lhe dis respeito eu aqui cheguei mto due.
is já estou com Algumas melhorias ainda q pouca

PSCR0521

Example 6: <salute> inside <opener>



[fig1]
Sr Dn Franco Anto de Picaza
Torrijos Fbro 16 / 95

Mui señor mio: Con muy
justa razon contemplo á U quexoso de mi fal-
ta de correspondencia, que está, ha
dependido

PSCR5532

Example 7: cross, <salute>, <letAddress> and <letDate> inside <opener>, and <salute> inside <p>

d. The titles and kinship relations are included inside the element <**signed**>:

Example 1:
<**signed**>El regidor Antonio Fernández del Valle</**signed**>

Example 2:
<**signed**>viuda de Antonio Solano</**signed**>

e. The drawing of a cross at the header of the letter is marked as <**salute**>. The abbreviation JMJ, equivalent to Jesus, Mary and Joseph, in the header of the letter is also marked with <**salute**>.

Example 1:
<**salute**><deco decoRef="fig1"/></**salute**>

Example 2:
<**salute**>JMJ</**salute**>

f. If the letter includes the name of the addressee at the end of the text, that information is treated as a <**salute**>, but with an attribute @**subcat** with the value “addressee”:

Example 1:

*Tu hermano que te estima,
Pedro Álvarez
Al señor Marcos Álvarez*

```
<salute>Tu hermano que te estima</salute>
<lb/>
<signed>Pedro Álvarez</signed>
<lb/>
<salute subcat="addressee">Al señor Marcos Álvarez</salute>
```

This same rule applies to those cases in which the place of destination is referred at the end of the letter.

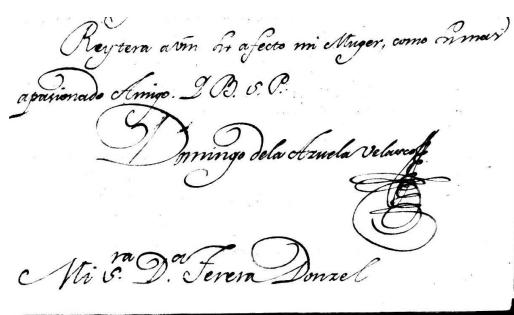
Example 2:

*Tu hermano que te estima
Pedro Álvarez
Cuenca, Al señor Marcos Álvarez*

```
<salute>Tu hermano que te estima</salute>
<lb/>
<signed>Pedro Álvarez</signed>
<lb/>
<letAddress subcat="addressee">Cuenca</letAddress>,
<salute subcat="addressee">Al señor Marcos Álvarez</salute>
```

- g. When the element **<salute subcat="addressee">** is graphically separated from the rest of the text (usually from the rest of the **<closer>**), this separation is marked using two elements **<lb/><lb/>** (with blank space before and after) before **<salute subcat="addressee">**.

```
<closer>
  <signed>Juan Hernandez</signed> <lb/><lb/>
  <salute subcat="addressee">Querido hermano Pedro Hernandez</salute>
</closer>
```



PS5132

Reytera a mi su afecto mi Muger, como su mas
apasionado Amigo Q B S P

Domingo de la Azuela Velasco

Mi sra Da Teresa Donzel

```
<closer>
  <signed>Domingo de la Azuela Velasco</signed> <lb/><lb/>
  <salute subcat="addressee">Mi sra Da Teresa Donzel</salute>
</closer>
```

- h. If the letter is signed by more than one autor, their names are declared in different elements **<signed>**:

```
<signed>Pedro Álvarez</signed>
<lb/>
<signed>Pablo Fernández</signed>
```

2.2.2. **<ps>**

- a. The element **<ps>** is used to transcribe the postscript of the letter. It contains an element **<p>** inside:

```
<ps>
  <p></p>
</ps>
```

- b. Any textual information that appears after the postscript and is not graphically detached is included inside this element:

```
<ps>
  <p>Y no te olvides de escribirnos. <signed>Agustín Pacheco</signed></p>
</ps>
```

If the textual content following the postscript is graphically detached, it is included inside the element **<closer>**:

```
<ps>
  <p>Y no te olvides de escribirnos.</p>
</ps>
<closer>
  <signed>Agustín Pacheco</signed>
</closer>
```

2.3. Full example of <text>

Example from letter PS6229¹⁶:

```
<text id="PS6229">
  <body><pb n="7r"/>
    <opener>
      <salute><deco decoRef="fig1"/></salute> <lb/>
      <letAddress>Madrid</letAddress> <letDate>Octubre 19 de 83</letDate>
    </opener>
    <p><salute>Esposa y querida mia</salute> me alegrar esta <lb/> te alle con la salu que yo para mi deseo manu<add
      hand="MS7" place="supralinear">e</add><lb subcat="false"/>la esta solo sirbe para decirte Ande as ido y <lb/> como
      te an recibido con buena cara o con mala <lb/> que todo lo quiero saber porque todavia no me<lb/> e muerto yo Anque
      estas trenita y dos leguas <lb/> desbiada de mi no te paresca te olvido que siempre <lb/> te tengo en el pensamiento
      ogala no te ubieras ido <lb/> que no me sucediera A mi lo que me sucede anque <lb/> no es nada ni lo sera si Ds||Dios
      quiere manuela todos <lb/> aquellos que saben que nos deben dinero se lo <lb/> cobraras sin dejar ninguno y
      principalmen<lb subcat="false"/>te Al ynfeliz de tu padre que te de todo lo que <lb/> nos debe sin perdonarle un ochabo
      y tambien <lb/> le sacaras todos los tratos que le quedamos ay <lb/> en esa su casa que no pido a Ds||Dios mas que
      cu<lb subcat="false"/>ando entre por ella se unda toda y lo sepulte <lb/> ay A ese infeliz Ai te remito la carta de tu <lb/>
      querida ermania y beras lo que te dice esa <unclear>sosa<lb subcat="false"/>manteles</unclear> tanbien te encargo le
      digas a tu<pb n="7v"/> padre el desdichao que mire no esta tratando <lb/> con musicos que trata con un onbre megor
      que <lb/> tu padre y que toda la musica de la catedral porque <lb/> como sepa que esta carta no se la les en su cara
      <lb/> sera cosa que yre yo A leerle el ultimo papel <lb/> de su bida tanbien en la segunda <gap reason="cancelled"
      extent="1 word"/> carta que es<lb subcat="false"/>pero si Ds||Dios te da salu me as de enbiar los recibos <lb/> que me
      tiene alla tu padre el bendito que son tres <lb/> los que tu no ynoras y tambien entre ellos me en<lb
      subcat="false"/>biaras el de tu tio Andre y le diras que si quiere le <lb/> enbie pluma y tinta para saber de su salu que
      no a sido <lb/> dino de enbiarme una mala carta Esto que te encargo <lb/> de los recibos de tu padre no lo eches en
      descuido porque <lb/> sabes bien mi genio que an de benir sin falta a la segunda <lb/> carta que me enbies Anque no
      me enbies un cuarto <lb/> que mas quiero no tener que comer que no que ese picaro <lb/> Aga burla de nosotros
      tanbien te digo te acuerdes como me de<lb subcat="false"/>gaste A que tu no eres tan falta de conocimiento como todo
      eso <lb/> y a tu padre le diras que el dinero del taller te lo ponga en la <lb/> mano Antes que Manuel Soler Pase a esa
      y delante <lb/> de el gastalo y luego cagate en el y el dinero que nos tiene. <lb/> mas y me escribiras si para bender la
      casa necesitas Poder<pb n="8r"/> mio para enbiartelo Al istante para que <lb/> cuanto Antes espaches con ello porque
      de <lb/> lo contrario me aras pasar a esa a perderme <lb/> de mi pretension te digo ba bien porque los dos <lb/> tios lo
      An cogido con mucho calor que esperan<lb subcat="false"/>cas en Ds||Dios tengo esperancas de que Ds||Dios me a
      degar <lb/> cagarne en todos; Tambien te encargo esto no <lb/> lo tomes a pechos porque ya sabes mi genio que no
      <lb/> puedo remediarlo y con esto daras memorias a <lb/> mi padre y ermania y al ama. y a andre y a la <lb/> celipa y
      las niñas y tu las tomaras a medida de <lb/> tu deseo quedo Pidiendo a Ds||Dios te Ge||Guardre <lb/> la bida
      ms||muchos as||años</p>
    </closer>
    <salute>quien te quiere y te <lb/> estima es tu esposo</salute> <lb/>
    <signed>Manuel Soler</signed>
  </closer>
</body>
</text>
```

¹⁶ The correspondence between this example and TEI-P5 is covered in section 4.2; there, the reference version of the text is the version edited in the TEITOK environment.

3. Transcription and editing conventions

In this section we offer an explanation of the underlying criteria for the XML markup of the texts. The adopted criteria include the expansion of all abbreviations (properly signalled), the standardization of word boundaries (except for Portuguese enclitic pronouns) and the distribution of the letters «i», «j», «u» y «v».

3.1. Standardization of punctuation

Certain punctuation marks used in the original manuscripts are not maintained in the transcription. That is the case of the following examples:

- a. The double slash used as paragraph mark (//) is transcribed as a full stop (.).

Original:	Esto es un ejemplo // Esto es un ejemplo
Transcription:	Esto es un ejemplo. Esto es un ejemplo

- b. The per cent sign is transcribed as colon (:).

Original:	Esto es un ejemplo % Esto es un ejemplo
Transcription:	Esto es un ejemplo: Esto es un ejemplo

- c. A segment of text written between slashes is transcribed between brackets.

Original:	/Esto es un ejemplo/
Transcription:	(Esto es un ejemplo)

- d. The equals sign used as paragraph mark (=) is transcribed as full stop (.)

Original:	Esto es un ejemplo = Esto es un ejemplo
Transcription:	Esto es un ejemplo. Esto es un ejemplo

- e. The sign "&" is transcribed in two different ways, according to its function in the text:

- If it has a coordinative function, it is transcribed as *e* (in Portuguese) or as *y* (in Spanish):

Original:	saludos & memorias para todos
Transcription:	saludos y memorias para todos

- If it represents the abbreviation of *etc*, it is transcribe as *e* and the abbreviation is expanded.(cf. 3.3.g)

3.2. Annotations appended to the text

When there are annotations by an external hand in the margins of the letter, that information is added in the context of the letter and not in the transcription. These annotations included in the context must be standardized for spelling.

3.3. Abbreviations

The markup of abbreviations is done without using XML elements, which allows for a faster transcription process. The strategy used consists on writing the original form of the abbreviation followed by its corresponding expanded form separating both forms by means of the sign ||.

Examples:

Example 1:
sr||señor

Example 2:
svor||servidor

When the XML file is uploaded into TEITOK and tokenized, all abbreviations are automatically converted into an XML format. The complete information for each token will be as follows:

Example 1:
`<tok id="w-1" fform="señor">sr</tok>`

Example 2:
`<tok id="w-1" fform="servidor">svor</tok>`

When dealing with abbreviations, the following general guidelines should be taken into account:

- a. All abbreviations inside **<body>** are expanded, but abbreviations inside **<envelope>** are not.
- b. Only the alphabetic content of each abbreviation is transcribed, without using any other symbol or punctuation, such as full stops, slashes or superscript characters, even if they appear in the original text. Therefore, even if the manuscript contains the form *no*, *nº* (with a superscript *o*), *no.* (with full stop after *o*) the transcription is always:

no||numero

- c. Abbreviations corresponding to more than one word (*sss*, *BSM*) have a sequence "Abbreviation||Expansion" for each word that integrates the abbreviation (for VMd, VMe, VSa, see below section i.):

Example 2:
B||Besa L||La M||Mano

Example 3:
S||Su S||Seguro S||Servidor

- d. Abbreviations containing regular ciphers (2o, 8bre), are marked and expanded:

Example 1:
2o||segundo

Example 2:
8bre||octubre

Secret ciphering systems are treated the same way. For example, if the hand replaces the vowels (a, e, i, o, u) by numbers (1, 2, 3, 4, 5) in the manuscript:

Original: 1m3g4
Transcription: 1m3g4||amigo

The type of codification used in the letter is explained in the context of the letter.

- e. Fully lexicalized abbreviations and monograms (etc, PD, PS, AD, XPO, XPTO, Lx, Lxa, IHS, JHS) are not marked with any other element, but the original punctuation and orthographical variation is respected (for VM, cf. m) below). A mixed case, that of the form *xrtobal* (Cristóbal), is treated as an abbreviation and, therefore, it is expanded:

xrtobal||xristobal

- f. All abbreviations are transcribed maintaining the spelling of the manuscript:

Example 1:
qdo||quando

Example 2:
affto||affecto

- g. The symbol & functioning as the abbreviation of etc can have three different forms: &, &a, &c.

Example 1: (&
e||et

Example 2: (&a)
ea||eta

Example 3: (&c)
ec|| etc

In all cases, the corresponding standardized form is "etc".

- h. Unknown abbreviations are not marked with any element, but a full stop is added. For example, an abbreviation *M* which refers to a person name that is unknown due to lack of information is transcribed as *M.* (*M* followed by full stop). The same rule applies to abbreviations by an illegible hand.
- i. Abbreviations used in the original with the intent to hide a person's identity are not marked with any element, even if that identity is known. In these cases, the abbreviation is kept, but with a full stop at the end, similarly to the case of unknown abbreviations (see the previous point).
- j. Abbreviations marked with diacritics or any symbol that encodes quantity. When a form *V̄I* is used to represent 6000 in Roman numbers and it is transcribed as follows:

VI||6000

The symbol used to represent the thousands that is used in some letters (see image below) is transcribed as *M* and the corresponding abbreviations is treated as follows:

6M||6000

Illustrated example:

*do efecto, por lo cual hoy, solo me hallo con .
5Dr. que puedo poner donde U. quiera, pues tenia
cerca de 99 y con tantos males en los crecidos*

PSCR5532

do efecto, por lo cual hoy, solo me hallo con
5M||5000 rs||reales que puedo poner donde U||Usted quiera, pues tenia
Cerca de 9M||9000 y con tantos males en los crecidos

If the implied quantity includes more ciphers, they should all be treated inside the same abbreviation:

5M500||5500

- k. Abbreviations of foreign words are expanded the same way as other abbreviations:

mm||madam

- l. The expansion of abbreviations should take into account the idiosyncrasies of the hand's spelling. Therefore, if the hand continuously uses "n" instead of "ñ", the abbreviation "as" for "años" should be transcribed as follows:

ES:
as||anos

In the same way, if throughout a Portuguese letter, the hand systematically uses the preposition “pera” instead of what became modern standard “para”, the abbreviation of that preposition should be expanded as follows:

PT:
pa || pera

- m. The abbreviations for “vuestra merced”, “vuestra señoría” (Spanish corpus), and “vossa mercê”, “vossa senhoria”, “vossa excelênciā” (Portuguese corpus), are not expanded. They are transliterated from the manuscript in their precise written form, respecting capitals wherever they occur. However, blank spaces are not transcribed.

Ejemplo 1:
ya no cango mas a vmd

Ejemplo 2:
quiera Ds||Dios que Vs se alle bien de salud

At the standardization level, these forms are spelled “VM”, “VSa”, “VExa” in the singular, and “VMs”, “VSAs”, “VExAs” in the plural. Other cases of forms of address, such as “VRa” (vuestra reverencia) or “Vlltma” (vuestra ilustrísima), are treated exactly like the rest of the abbreviations formed by two or more words (cf. 3.3.c.).

3.4. Beginning of folio **<pb @n/>**

- a. The beginning of the folio is marked with the element **<pb @n/>**, which appears immediately after the element **<body>**. The following page breaks are marked when there is a folio change in the original manuscript.

```
<body><pb n="16r"/>
  <opener/>
    <p>Texto<pb n="16v"/> texto<pb n="17r"/> texto</p>
  <closer/>
</body>
```

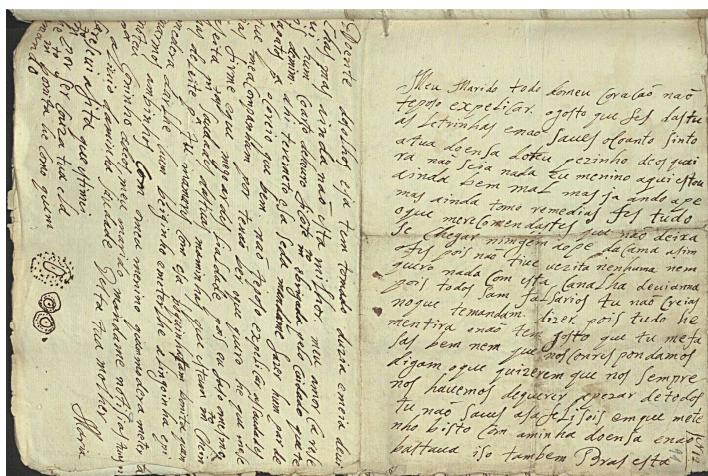
- b. The attribute **@n** contains the number of the corresponding folio, according to the original numeration, followed by the letters r (recto) or v (verso).

```
<pb n="16r"/>
```

If the number of the folio is conjectured, it appears within square brackets:

```
<pb n="["16]r"/>
```

- c. Bare in mind that the number of folios must always match the exact foliation of the folio containing the letter within the documental unit. This means that the numeration of each folio is independent from any possible folding the paper might have suffered before being attached to the documental unit. Paper folding is only taken into account while physically describing the manuscript inside element <**support**> (cf. 1.1.3.3.2.).
- d. Regardless of what is said in the previous point, a letter that was archived as one single folio may present internal folds which represent pagination marks created during the writing process. In these cases, it is necessary to use an element <**pb**> to mark the page breaks, however, repeating the same value for the **attribute @n**, in order to respect the foliation of the documental unit.



CARDS2708

The picture above presents several folds. However, only the fold that divides the letter in two corresponds to a page break created during the writing process (all other folds were done after the writing process). The correct transcription of the letter should be as follows:

```

<pb n="91r"/>
<p>Meu Marido todo do meu CoraÇao ....
.....
bastava iso tambem Bras esta <pb n="91r"/>
Doente dos olhos e ja tem tomado duzia ...
</p>

```

Furthermore, bare in mind that, in this case, we don't mark a vertical orientation of the text, since there isn't an orientation change inside the same <**pb**> (cf. 3.6 and 3.7.)

3.5. Beginning and end of paragraph <p></p>

Paragraphs begin with an element <p> and they are closed with and element </p>. There are as many paragraph changes as paragraphs in the manuscript. It is considered a paragraph change whenever there is a graphic mark in the original, being it one or more blank lines or an interrupted line:

Esto es un ejemplo, esto es un ejemplo,
esto es un ejemplo, esto es un ejemplo.

Esto es un ejemplo, esto es un ejemplo,
esto es un ejemplo.
Esto es un ejemplo, esto es un ejemplo,
esto es un ejemplo, esto es un ejemplo.

<p>Esto es un ejemplo, esto es un ejemplo,
esto es un ejemplo, esto es un ejemplo.</p>

<p>Esto es un ejemplo, esto es un ejemplo,
esto es un ejemplo</p>.

<p>Esto es un ejemplo, esto es un ejemplo,
esto es un ejemplo, esto es un ejemplo.</p>

3.6. Beginning of line: <lb/>

The beginning of a line is marked with an element <lb/>. There are as many line beginnings as the number os lines in the manuscript. There are five possibilities:

Beginning of a line after a graphic element: <pb/>, <p>, <opener> and <closer>. In these cases, there is no need to mark the beginning of the line with an <lb/> (this is an automatic process when the letter is tokenized)

```
<body><pb n="16r">
  <opener><salute>Querido amigo</salute></opener>
  <p>Te escribo esta carta para decirte... </p>
</body>
```

- Beginning of a line between two independent words. An element <lb/> is used with a blank space before and after:

Te escribo para <lb/> decirte...

- Beginning of a line in the middle of a word. In this case, the attribute @n="false" is used and the element <lb/> is not preceded nor followed by a blank space:

Te escribo para de<lb subcat="false"/>cirte...

- Beginning of a line with vertical orientation between independent words. In this case, the attribute line_rend@="verticalText" is used. The element <lb/> is preceded and followed by a blank space:

Te escribo para <lb line_rend="verticalText"/> decirte...

- d. Beginning of a line with vertical orientation in the middle of a word. In this case, the attributes **line_rend@="verticalText"** and **n@="false"** are used. The element **<lb/>** is not preceded nor followed by a blank space:

Te escribo para de<lb line_rend="verticalText" subcat="false"/>cirte...

The cases described above can coincide with the beginning of a folio. The following possibilities can occur:

- e. Beginning of a folio between two independent words. The element **<lb/>** is not marked (these cases are treated automatically when the letter is tokenized):

Te escribo para<pb n="28r"/> decirte

- f. Beginning of a folio in the middle of a word. The element **<lb/>** is not preceded nor followed by a blank space:

Te escribo para de<pb n="28r"/><lb subcat="false"/>cirte

The line beginnings that occur between two different elements of a formulaic part (**<opener>**, **<closer>**, **<salute>**, **<letAddress>**, **<letDate>**, **<signed>**) are encoded outside these elements (cf. 2.2.1.a):

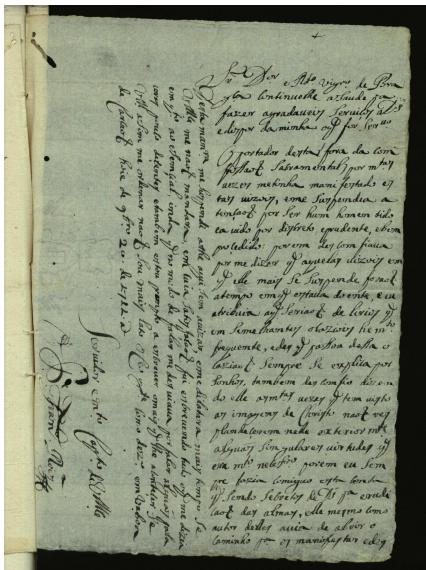
```
<closer>
    <salute>Tu hermano que te estima</salute> <lb/> <signed>Pepe</signed>
</closer>
```

About the use of **<lb/><lb/>** before an element **<salute>** (cf. 2.2.1.g).

3.7. Vertical text: **<lb @line_rend/>**

To transcribe a sequence of text that was written with a vertical orientation, we use the element **@line_rend="verticalText"** (cf. 3.6.).

Vertical text is only encoded when text orientation suffers a shift inside the same folio, that is to say, inside the one same unit delimited by **<pb/>**.



CARDS2107

Sr Dor e Rdo Vigro de Bra <lb/>

... <lb/>

... <lb/>

... <lb/>

caminho pa os manifestar e des <lb line rend="verticalText"/> Desta manra me suspendi athe aqui sem avizar, e me dilatara mais tempo se <lb line rend="verticalText"/> VMce me não mandara, em cuja satisfação fui escrevendo tudo o q me dizia <lb line rend="verticalText"/> em qto ao esemcial, inda q no modo de falar me desviava por falar alguãs pala <lb line rend="verticalText" subcat="false"/> vrás pouco deçentes e tambem estou pronto a escrever o mais q lhe acontecer se ...

Compare these image (CARDS2170), which contains vertical text, with the previous image (CARDS2708), which doesn't have vertical text.

3.8. Corrected text:

- The element is used to mark a sequence of text that has been stricken-through by the hand, but whose content is possible to understand. The element is always followed by the attribute @hand, in which is included the reference (using the corresponding initials) of the responsible for the erasure (the author, the scribe):

Example 1 (character stricken-through by Agustín Pacheco)

Original: escribí a Paedro Cortés

Transcription: escribí a P<del hand="AP1">adro Cortés

Example 2 (word stricken-through by an annotator)

Original: la mujer del comerciante

Transcription: la <del hand="annotator">mujer del comerciante

Illustrated example:

PS9026

<del hand="VF3">la gallega te respondi no tube lugar para ...

- b. The element **** may apply to parts of words (example 1 above), whole words (example 2 above) or sequences of words in cases such as the one below:

Original: escribí a mi primo Juan Pedro Cortés
 Transcripción: escribí a <del hand="AP1">mi primo Juan Pedro Cortés

- c. Whenever a word is replaced by another word outside the line, the correction is transcribed inside the element **<add>** (cf. 3.9). The element **<add>** contains an attribute **@hand**, which includes the reference of the participant responsible for the correction and an attribute **@place**, which contains information on the spot of the layout where the correction was added. For instance, a replacement of “Juan” by “Pedro” made by Agustín Pacheco in an upper line would be transcribed as follows:

Original: escribí a Juan Pedro Cortés
 Transcription: escribí a <del hand="AP1">Juan <add hand="AP1" place="supralinear">Pedro</add> Cortés

The element **<add>** is only used when its content (the added element) was written outside the line level. If the deletion and the addition are written in the line sequence, the element **<add>** is not used. In that case, the added word is written on the right of the deleted word:

Original: escribí a Juan Pedro Cortés
 Transcripción: escribí a <del hand="AP1">Juan Pedro Cortés

- d. If the deletion is impossible to understand, the element **** is not used. In these cases, the element **<gap>** is used instead. This element contains two attributes: **@reason** (the reason for omitting the text, which is always **@reason="cancelled"** when the text is stricken-through) and **@extent** (the conjectured extension of the omitted text):

Original: la —— del comerciante
 Transcription: la <gap reason="cancelled" extent="1 word"/> del comerciante

3.9. Added text: **<add>**

- a. The element **<add>** is used to mark a sequence of text that has been added outside the line level. It contains two attributes: **@hand** (person responsible for the added text) and **@place** (place where the text was added).

Original: escribí a Pedro Cortés
 Transcripción: escribí a <add hand="AP1" place="supralinear">Pedro</add> Cortés

The possible values for the element @place are:

place="supralinear"
place="underlinear"
place="margin-top"
place="margin-left"
place="margin-right"
place="margin-bottom"
place="back"

Illustrated example:

Luego q! los Yntervalos de locura te dexen en libertad
devorar seguiran con toda prudencia y pureza mis caras

PSCR5502

Luego qe los yntervalos de **<add hand="scribe" place="supralinear">**tu**</add>** locura te dexen en libertad

- b. The element **<add>** may apply to letters, whole words or sequences of words, just like the element **** (cf. 3.8.):

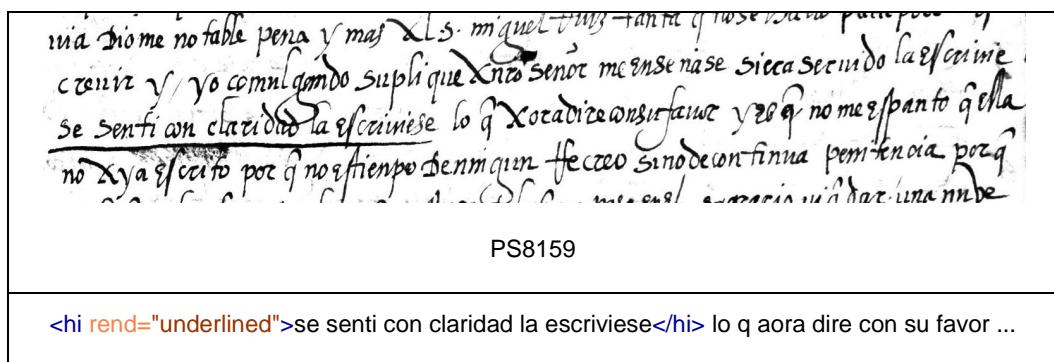
Original: escribí a mi primo Pedro Cortés
Transcripción: escribí a **<add hand="AP1" place="supralinear">**mi primo Pedro**</add>**
Cortés

3.10. Underlined text: **<hi>**

- a. A sequence of text underlined by the original hand is marked with the element **<hi>**. This element receives the attribute @rend="underlined".

Original: Te escribo esta carta para decirte...
Transcription: Te escribo **<hi rend="underlined">**esta carta**</hi>** para decirte...

Illustrated example:



PS8159

<hi rend="underlined">se senti con claridad la escriviese**</hi>** lo q aora dire con su favor ...

- b. If the text was underlined by an annotator, including the addressee, the procedure should be the following:

```
<hi rend="underlined" subcat="annotator">texto subrayado</hi>
```

In this case, the existence of an annotator must be previously declared in the element **<letHeading>** (cf. 1.1.3.2.1.g):

```
<respStmt><resp subcat="annotator" id="annotator"/></respStmt>
```

- c. Element **<hi>** applies to letters, whole words or sequences of words, according to each case.

3.11. Omitted text: **<gap/>**

- a. If a sequence of text is not possible to read because it is stricken-through, the characters are hidden or the support is damaged, we use the element **<gap/>**. This element contains, at least, two attributes: **@reason** (the reason for the text omission) and **@extent** (the approximate extension of the omitted text: 1 word, 3 words, 2 lines, etc.):

Example 1: sequence of three illegible words due to the a difficult hand
`<gap reason="illegible" extent="3 words"/>`

Example 2: sequence of two illegible words due to an accident in the physical support
`<gap reason="damage" extent="2 words"/>`

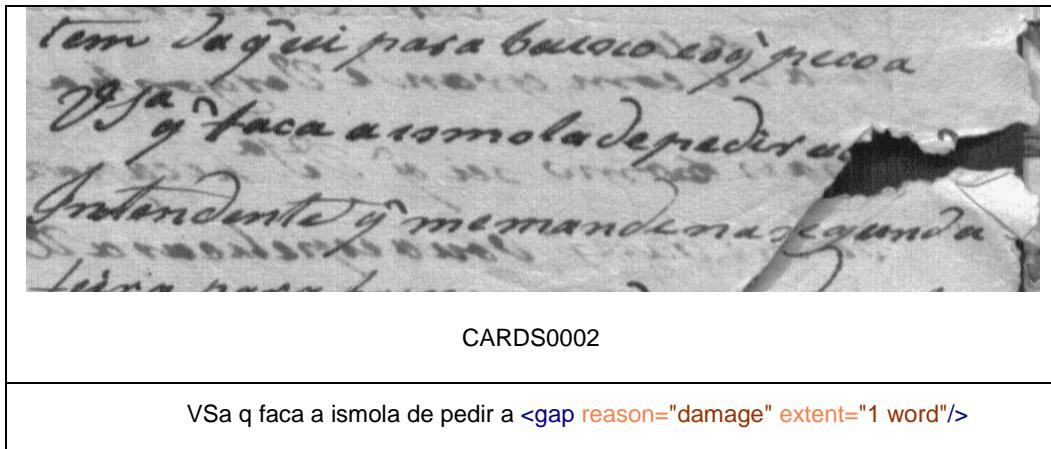
Example 3: sequence of one word hidden by the stitching of the gathering
`<gap reason="hidden" extent="1 word"/>`

The possible values for the attribute **@reason** are:

`reason="damage"`
`reason="hidden"`

```
reason="cancelled"  
reason="illegible"  
reason="otherLetter"
```

Illustrated example:

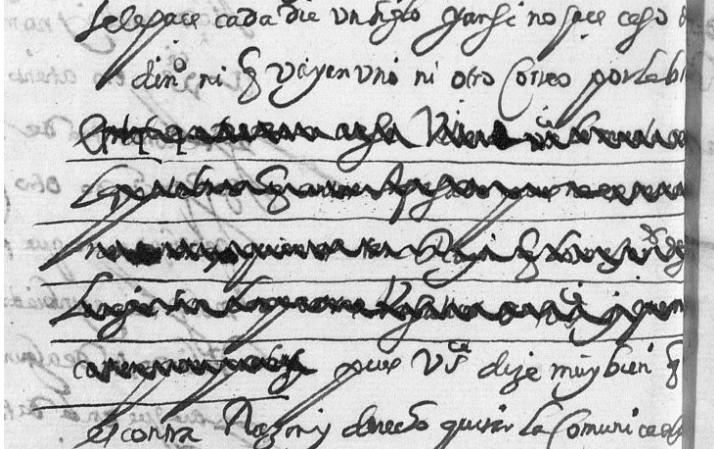


- b. The element **<gap/>** is not used with sequences smaller than a word. If there are illegible characters within a word, the element **<gap/>** replaces the whole word with the attribute **@extent="1 word"**.
- c. The element **<gap>** can't be used with sequences larger than a line, except in cases of a letter within a letter (cf. 3.17.a.). If there is an illegible sequence that includes several lines, the procedure should be as follows.

texto texto texto [rotura]
[rotura]
[rotura]
[rotura] texto texto texto

texto texto texto <gap reason="damage" extent="X words"/>
<gap reason="damage" extent="1 line"/>
<gap reason="damage" extent="1 line"/>
<gap reason="damage" extent="X words"/> texto texto texto

Illustrated example:



PS5021

dino ni q vayan uno ni otro correo por la brevd
<gap reason="cancelled" extent="1 line"/>
<gap reason="cancelled" extent="1 line"/>
<gap reason="cancelled" extent="1 line"/>
<gap reason="cancelled" extent="1 line"/>
<gap reason="cancelled" extent="2 words"/> pues vsia dize muy bien q

- d. If an illegible word is interrupted by a line break, the element **<gap>** replaces the whole word, but it is only marked after the line break:

a causa de una [rotura]
tión de mi hermano

a causa de una <lb/> <gap reason="damage" extent="1word"/> de mi hermano

- e. The attribute **@reason="otherLetter"** is used when a letter contains another letter in its text (cf. 3.17.a.).

3.12. Supplied text: **<supplied>**

- a. If it is necessary to conjecture a sequence of text, the element **<supplied>** is used. This element contains two attributes: **@resp** (for the initials of the person responsible for the conjecture) and **@reason** (for the reason that motivated the conjecture).

Example 1: conjecture by Gael Vaamonde because of an illegible hand
<supplied resp="GV" reason="illegible">amigo</supplied>

Example 2: conjecture by Rita Marquihdas because of a damaged support
<supplied resp="RM" reason="damage">amigo</supplied>

Example 3: conjecture by Laura Martínez Martín because a word got hidden inside the gathering spine
`<supplied resp="LMM" reason="hidden">amigo</supplied>`

Example 4: conjecture by Elisa García Prieto because one or several words were cancelled

```
<supplied resp="EGP" reason="cancelled">amigo</supplied>
```

The possible values for the attribute @reason are:

reason="damage"
reason="hidden"
reason="abandoned"
reason="illegible"
reason="code"
reason="cancelled"

Illustrated example:



- b. The element **<supplied>** is only used for whole words. Even if only a part of the word is conjectured, the whole is included in the element **<supplied>**. If more than one word is conjectured, an element **<supplied>** will be open for each conjectured word.
 - c. Conjecturing abandoned words. The attribute **@reason="abandoned"** inside **<supplied>** is used to conjecture possible erasures. Certain letters or words may have been written by the hand in a first moment and then abandoned, in a revision campaign, without being stricken-through. Therefore, what is being conjectured is the possible deletion:

Original: recibimos mos
Transcription: recibimos <supplied resp="GV" reason="abandoned"><del hand="AP1">mos<supplied>

This strategy is only used if it is possible to conjecture an oversight on behalf of the author or scribe. It is not used when there is a repetition of a word at the end of a folio and at the beginning of the next folio (in the cases of catchwords), because, in these cases, the repetition is intentional.

3.13. Unclear text (*locus desperatus*): <unclear>

- The element <**unclear**> is used whenever a word cannot be read with certainty because the text, although legible, is difficult to understand:

<unclear>lesxdon</unclear>

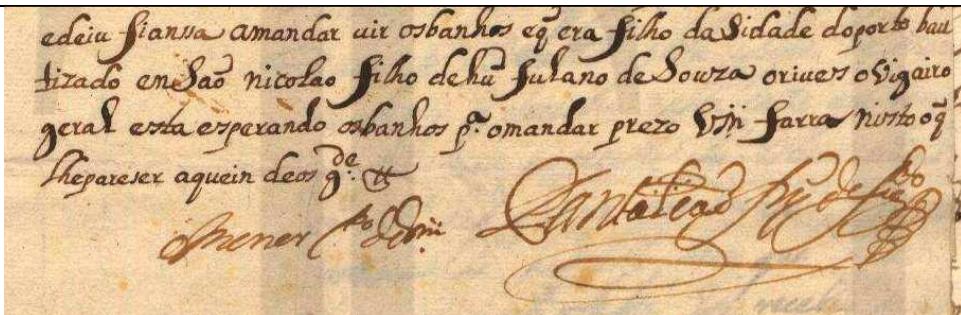
- The element <**unclear**> is used for whole words or groups of words, but not with parts of words.
- The element <**unclear**> is not used if the undecipherable text is an abbreviation. In this case, the letters that form the abbreviation are transcribed followed by full stop (cf. 3.3.h).

3.14. Text from another hand: <handShift/>

- The sequences of text written by a new hand are encoded with the element <**handShift**>. This element appears at the beginning of the text written by a new hand:

<signed><handShift/>Maria Josefa</signed>

Illustrated example:



CARDS1043

<handShift/><salute>menor Cro de Vm</salute> <signed>Pantaleão frco de fgdo</signed>

- The element <**handShift**> indicates that, from that moment on, in the manuscript, the text started to be written by another hand. Therefore, if, after that, there is text written by the first hand (the original hand), a new element <**handShift**> must be used.
- If there is an ID for the new hand, a new element <**author**>, corresponding to the new hand, must be added inside the element <**letHeading**>. In those cases, in

the transcription, the element **<handShift/>** will include an attribute **@resp** with the initials of the responsible for text written by the new hand :

```
<handShift resp="AP1"/>
```

3.15. Discourse organization: <s formula>

- a. The encoding of the letter's discourse organization is done using three possible discursive parts: harangue, narration and peroration. This information is encoded in TEITOK once the text has been automatically split into sentences (cf. paragraph 4).
- b. The XML element used in TEITOK to split sentences is **<s>**. Discursive segments are encoded inside each element **<s>** with an attribute **@formula**. The value of this attribute specifies the discursive type:

```
<seg subcat="harangue">sentence marked as harangue</seg>
<seg subcat="narration">sentence marked as narration</seg>
<seg subcat="peroration">sentence marked as peroration</seg>
```

- c. The element **<s>** combined with its attribute **@formula** are also used to encode formulaic parts in the body of the letter, that is **<salute>**, **<signed>**, **<letAddress>** and **<letDate>** **<p>**. This way, the formulaic parts in the beginning of **<p>** are included in an element **<s>** with the attribute **@formula="opener"**; the formulaic parts at the end of **<p>** are included in an element **<s>** with the attribute **@formula="closer"**:

Example 1 (CARDS2701):

```
<s formula="opener"><salute>Meu marido e todo o meu bem</salute></s>
```

Example 2 (PS8098):

```
<s formula="closer"><letAddress>San Bernardino</letAddress>, <letDate>hoy domingo 1 de marzo de 1732</letDate></s>
```

- d. *Harangue* and *peroration* are encoded manually. *Narration* is marked automatically (i.e. everything that isn't *harangue* nor *peroration*). Both **<s formula="opener">** and **<s formula="closer">** are encoded automatically, although manual revision may be required when there is no matching at the beginning or end of **<p>**.

Splitting text into sentences only occurs inside element **<p></p>**. There is no sentence splitting inside the elements **<opener>** and **<closer>**. The formulaic parts included inside **<opener>** and **<closer>** are encoded using the elements **<salute>**, **<letAddress>**, **<letDate>** and **<signed>**. (cf. 2.2.1.)

3.16. Postmark: <postmark>

- a. The element <postmark> is used to encode text imprinted on the envelope. Inside <postmark> we use the element <placeName> if the postmark contains the name of a place, or <date> if the postmark refers to a date:

```
<postmark><placeName>Lisboa</placeName></postmark>
```

- b. The element <postmark> can only be used in the envelope, inside the element <envPart>:

```
<envelope type="superscript">
  <envPart>
    <postmark><placeName>Lisboa</placeName></postmark>
    <address type="receiver">
      <addrLine>A mi hermano Pedro</addrLine>
    </address>
  </envPart>
</envelope>
```

- c. It is mandatory that the element <postmark> contains a <placeName> and a <date>. If we have no hint on such information, we still use [S.I.], initials of “Sin localización” (no place):

```
<postmark><placeName> [S.I.]</placeName></postmark>
```

3.17. Other cases:

- a. Letter inside a letter:

- Copy X of a letter of which there is no original manuscript inside a letter Y. Two XML files are created, one for each letter (X and Y). In the transcription of the letter Y, the following information is added right at the place where the Copy X starts:

```
<gap reason="otherLetter" extent="X lines"/>
```

The attribute @extent contains the number of lines occupied by X within Y. The case is properly explained in the context.

- Copy Z of an original inside a letter Y. Two XML files are created, one for X and the other for Y. In the transcription of the letter Y the following information is added right at the place where the copy Z starts:

```
<gap reason="otherLetter" extent=X lines""/>
```

The attribute @extent contains the number of lines occupied by Z within Y. In the contexts of letters X and Y the case is properly

explained and the text of the copy Z is transcribed, in a conservative way.

- b. Draft and final version of a letter. Only an XML file is created, for the final version. The text of the draft is transcribed in the context, in a conservative way, and an explanation of the case is also added.

4. XML changes in TEITOK

4.1. General explanation

The digital edition workflow in *P.S. Post Scriptum* is composed by two main tasks: the text encoding using XML and the linguistic annotation using TEITOK. The first task includes the transcription, the scholarly edition of the manuscript, as well as full encoding of the header (`<teiHeader>`) and the encoding of discourse parts (harangue, narration and peroration). This task results in an XML file with textual and extra-textual information of a handwritten letter.

This XML file is then imported into the TEITOK interface where manual corrections and automatic annotations are added. The text is tokenized, the spelling is standardized, it receives lemmatization and morphosyntactic annotation¹⁷. The text included in `<p></p>` is split into sentences and the discursive parts of the text are marked (harangue, narration and peroration). The changes applied to the XML file are the following:

- Tokenization: when the text is tokenized, each token (words or punctuation marks) is automatically included inside an element `<tok>`, which receives a unique identification.
- Automatic changes: when tokenizing the text, there are three automatic changes that must be applied: one of the changes concerns abbreviations, the second is related to the element of beginning of line (cf. 3.4) and the other one is related to the beginning of a folio (cf. 3.6):
 - a) Conversion of abbreviations from a previous format (eg. sr||señor) to an XML format (ex. `<tok id="w-10" fform="señor">sr</tok>`)
 - b) Insertion of a `<lb/>` after every graphic element: `<pb>`, `<p>`, `<opener>` and `<closer>`.
 - c) The insertion of an attribute `@facs` inside the element `<pb>`. This attribute relates each folio to the corresponding image, allowing for the simultaneous visualization of text and image.
- Split into sentences: the text inside `<p></p>` is automatically split into sentences (`<s></s>`). Each sentence is automatically limited by strong punctuation marks, that is, full stop (.), exclamation mark (!), question mark (?). An unique identification is automatically given to each element `<s>`.

¹⁷ The lemmatization and morphosyntactic tasks are still in a trial period.

Optionally, each element **<s>** can adopt an attribute **@formula** which gives information on the discursive type or formulaic part, accordingly.

- Discursive organization and formulaic parts: the discursive organization of the text is done after the automatic splitting into sentences (**<s></s>**). Three main types of discursive parts are considered (cf. 3.15.a):

```
<s id="s-1" formula="harangue">sentence marked as harangue</s>
<s id="s-1" formula="narration">sentence marked as narration</s>
<s id="s-1" formula="peroration">sentence marked as peroration </s>
```

Sentences that include a formulaic part inside **<p>** (i.e. not graphically detached from the body of the letter) are also marked with the attribute **@formula**, which acquires the values “opener” and “closer”, accordingly.

```
<s id="s-1" formula="opener">sentence marked as opener</s>
<s id="s-1" formula="closer">sentence marked as closer</s>
```

Example:

Example 1 (CARDS2701):

```
<s formula="opener"><salute>Meu marido e todo o meu bem</salute></s>
```

Example 2 (PS8098):

```
<s formula="closer"><letAddress>San Bernardino</letAddress>, <letDate>hoy domingo 1 de marzo de 1732</letDate></s>
```

- Each element **<tok>** can have a series of attributes to encode the different levels of linguistic annotation:

form=" "	original form of the word ¹⁸
fform=" "	expanded form (in the case of abbreviations), free form (in the case of contractions)
dform=" "	dialectal or non-standard form
nform=" "	normalized form
pos=" "	morphosyntactic tag
lemma=" "	lemma
ltags=" "	linguistic keywords

- Inside **<tok>**, each contraction or enclisis results in a division into **<d tok>** elements, which receive the same levels of annotation of **<tok>**, using the attributes **@form**, **@pos**, **@mfs** and **@lemma** and the optional attributes **@fform** (meaning free form of a contracted element), **@nform** and **@dform**.

¹⁸ For each token, the attribute **@form** is filled only when the manuscript presents any XML element in the middle of the word: line break, erasure or addition.

4.2. Example

Fragment from letter PS6229 (cf. 2.3):

```
<text id="PS6229">
  <body><pb n="7r" facs="PS6229_1.JPG" admin="1" id="e-1"/>
    <opener>
      <salute><deco decoRef="fig1" id="e-2"/></salute> <lb id="e-3"/>
      <letAddress><tok id="w-1" pos="N" mfs="NP00000" lemma="madrid">Madrid</tok><tok nform="," id="w-2" pos="F" mfs="Fc" lemma=","><ee/></tok></letAddress>
      <letDate><tok id="w-3" nform="octubre" pos="N" mfs="NCMS000" lemma="octubre">Octubre</tok> <tok id="w-4" pos="Z" mfs="Z" lemma="19">19</tok> <tok id="w-5" mfs="SPS00" lemma="de">de</tok> <tok id="w-6" pos="Z" mfs="Z" lemma="83">83</tok></letDate>
    </opener>
    <p id="p-1"><s id="s-2" formula="opener"><salute><tok id="w-7" pos="N" mfs="NCFS000" lemma="esposo">Esposa</tok> <tok id="w-8" pos="C" mfs="CC" lemma="y">y</tok> <tok id="w-9" pos="N" mfs="NCFS000" lemma="querido">querida</tok> <tok id="w-10" nform="mía" pos="P" mfs="PX1FS0S0" lemma="mío">mía</tok><tok nform="." id="w-11" pos="F" mfs="Fp" lemma=". "><ee/></tok></salute></s> <s id="s-3" formula="harangue"><tok id="w-12" nform="Me" pos="P" mfs="PP1CS000" lemma="me">me</tok> <tok id="w-13" nform="alegraré" pos="V" mfs="VMIF1S0" lemma="alegrar">alegrar</tok> <tok id="w-14" nform="ésta" pos="P" mfs="PD0FS000" lemma="este">esta</tok> <lb id="e-4"/><tok id="w-15" pos="P" mfs="PP2CS000" lemma="te">te</tok> <tok id="w-16" nform="halle" pos="V" mfs="VMPS3S0" lemma="hallar">alle</tok> <tok id="w-17" mfs="SPS00" lemma="con">con</tok> <tok id="w-18" pos="D" mfs="DA0FS0" lemma="el">la</tok> <tok id="w-19" nform="salud" itags="complex_syllable" pos="N" mfs="NCFS000" lemma="salud">salu</tok> <tok id="w-20" pos="P" mfs="PROCN000" lemma="que">que</tok> <tok id="w-21" pos="P" mfs="PP1CSN00" lemma="yo">yo</tok> <tok id="w-22" mfs="SPS00" lemma="para">para</tok> <tok id="w-23" nform="mí" pos="P" mfs="PP1CS00" lemma="mí">mi</tok> <tok id="w-24" pos="V" mfs="VMIP1S0" lemma="desear">deseo</tok><tok nform="." id="w-25" pos="F" mfs="Fp" lemma=". "><ee/></tok></s> <s id="s-4"><tok id="w-26" form="manuela" nform="Manuela" pos="N" mfs="NP00000" lemma="manuela">manu<add hand="MS7" place="supralinear">>e</add><lb id="e-5"/>la</tok><tok nform="," id="w-27" pos="F" mfs="Fc" lemma=". "><ee/></tok> <tok id="w-28" nform="ésta" pos="P" mfs="PD0FS000" lemma="este">esta</tok> <tok id="w-29" nform="sólo" pos="R" mfs="RG" lemma="sólo">solo</tok> <tok id="w-30" nform="sirve" pos="V" mfs="VMIP3S0" lemma="servir">srirbe</tok> <tok id="w-31" mfs="SPS00" lemma="para">para</tok> <tok id="w-32" form="decirte">>decirte</dtok id="d-32-1" pos="V" mfs="VMN0000" lemma="decir"/><dtok id="d-32-2" pos="P" mfs="PP2CS000" lemma="te"></tok> <tok id="w-33" nform="ande" dform="ande" pos="P" mfs="PT000000" lemma="donde">Ande</tok> <tok id="w-34" nform="has" pos="V" mfs="VAIP2S0" lemma="haber">as</tok> <tok id="w-35" pos="V" mfs="VMP0000" lemma="ir">ido</tok> <tok id="w-36" pos="C" mfs="CC" lemma="y">y</tok> <lb id="e-6"/> <tok id="w-37" nform="cómo" pos="P" mfs="PT000000" lemma="cómo">como</tok> <tok id="w-38" pos="P" mfs="PP2CS000" lemma="te">te</tok> <tok id="w-39" nform="han" pos="V" mfs="VAIP3PO" lemma="haber">an</tok> <tok id="w-40" pos="V" mfs="VMP0000" lemma="recibir">recibido</tok><tok nform="," id="w-41" pos="F" mfs="Fc" lemma=". "><ee/></tok> <tok id="w-42" mfs="SPS00" lemma="con">con</tok> <tok id="w-43" pos="A" mfs="AQ0FS0" lemma="bueno">bueno</tok> <tok id="w-44" pos="N" mfs="NCFS000" lemma="cara">cara</tok> <tok id="w-45" pos="C" mfs="CC" lemma="o">o</tok> <tok id="w-46" mfs="SPS00" lemma="con">con</tok> <tok id="w-47" pos="A" mfs="AQ0FS0" lemma="malo">mala</tok><tok nform="." id="w-48" pos="F" mfs="Fp" lemma=". "><ee/></tok></s> <s id="s-5"><lb id="e-7"/> <tok id="w-49" nform="Que" pos="C" mfs="CS" lemma="que">que</tok> <tok id="w-50" pos="P" mfs="PI0MS000" lemma="todo">todo</tok> <tok id="w-51" pos="P" mfs="PP3CNA00" lemma="lo">lo</tok> <tok id="w-52" pos="V" mfs="VMIP1S0" lemma="querer">quiero</tok> <tok id="w-53" pos="V" mfs="VMN0000" lemma="saber">saber</tok> <tok id="w-54" pos="C" mfs="CS" lemma="porque">porque</tok> <tok id="w-55" nform="todavía" pos="R" mfs="RG" lemma="todavía">todabia</tok> <tok id="w-56" pos="R" mfs="RN" lemma="no">no</tok> <tok id="w-57" pos="P" mfs="PP1CS000" lemma="me">me</tok> <lb id="e-8"/> <tok id="w-58" nform="he" pos="V" mfs="VAIP1S0" lemma="haber">e</tok> <tok id="w-59" pos="V" mfs="VMP0000" lemma="morir">muerto</tok> <tok id="w-60" pos="P" mfs="PP1CSN00" lemma="yo">yo</tok><tok nform="." id="w-61" pos="F" mfs="Fp" lemma=". "><ee/></tok></s> <s id="s-6"><tok id="w-62" nform="Aunque" itags="diphthong_and_hiatus" pos="C" mfs="CC" lemma="aunque">Anque</tok> <tok id="w-63" nform="estás" pos="V" mfs="VAIP2S0" lemma="estar">estas</tok> <tok id="w-64" nform="treinta" itags="complex_syllable" pos="Z" mfs="Z" lemma="30">trenita</tok> <tok id="w-65" pos="C" mfs="CC" lemma="y">y</tok> <tok id="w-66" pos="Z" mfs="Z" lemma="2">dos</tok> <tok id="w-67" pos="N" mfs="NCFP000" lemma="legua">leguas</tok> <lb id="e-9"/> <tok id="w-68" nform="desviada" pos="V" mfs="VMP00SF" lemma="desviar">desbiada</tok> <tok id="w-69" mfs="SPS00" lemma="de">de</tok> <tok id="w-70" nform="mí" pos="P" mfs="PP1CS00" lemma="mí">mi</tok><tok nform="," id="w-71" pos="F" mfs="Fc" lemma=". "><ee/></tok>
```

TEI-P5

Initial fragment from example 2.3. and 4.2.¹⁹

```

<text xml:id="PS6229">
  <body><pb n="7r" facs="PS6229_1.JPG"/>
    <opener>
      <salute><seg type="sign">cruz</seg></salute> <lb break="yes"/>
      <placeName><w lemma="madrid">Madrid</fs type="pos"><f name="N"><symbol value="NP00000"/></f></fs></w><pc change=". " type="F" subtype="Fc"/></placeName>
      <date><w lemma="octubre"><choice><orig>Otubre</orig><reg>octubre</reg></choice><fs type="pos"><f name="N"><symbol value="NCMS000"/></f></fs></w> <w lemma="19">19<fs type="pos"><f name="Z"><symbol value="Z"/></f></fs></w> <w lemma="de">de<fs type="pos"><f name="S"><symbol value="SPS00"/></f></fs></w> <w lemma="83">83<fs type="pos"><f name="Z"><symbol value="Z"/></f></fs></w></date>
    </opener>
    <p><s><seg type="salute"><w lemma="esposo">Esposa</fs type="pos"><f name="N"><symbol value="NCFS000"/></f></fs></w> <w lemma="y">y<fs type="pos"><f name="C"><symbol value="CC"/></f></fs></w>
      <w lemma="querido">querida<fs type="pos"><f name="N"><symbol value="NCFS000"/></f></fs></w> <w lemma="mío">mío</fs type="pos"><f name="P"><symbol value="PX1FS0S0"/></f></fs></w><pc change=". " type="F" subtype="Fp"/></seg></s><seg type="harangue"><s> <w lemma="me"><choice><orig>mía</orig><reg>mía</reg></choice><fs type="pos"><f name="P"><symbol value="PP1CS000"/></f></fs></w> <w lemma="alegrar"><choice><orig>alegrar</orig><reg>alegrar</reg></choice><fs type="pos"><f name="V"><symbol value="VMIF1S0"/></f></fs></w> <w lemma="este"><choice><orig>esta</orig><reg>ésta</reg></choice><fs type="pos"><f name="P"><symbol value="PD0FS000"/></f></fs></w> <lb break="yes"/> <w lemma="te">te<fs type="pos"><f name="P"><symbol value="PP2CS000"/></f></fs></w> <w lemma="hallar"><choice><orig>alle</orig><reg>halle</reg></choice><fs type="pos"><f name="V"><symbol value="VMPS3S0"/></f></fs></w> <w lemma="con">con<fs type="pos"><f name="S"><symbol value="SPS00"/></f></fs></w> <w lemma="el">la<fs type="pos"><f name="D"><symbol value="DA0FS0"/></f></fs></w> <w><choice><orig>salu</orig><reg>salud</reg></choice><fs type="pos"><f name="N"><symbol value="NCFS000"/></f></fs></w> <note type="linguisticNote" subtype="complex_syllable"/></w> <w lemma="que">que<fs type="pos"><f name="P"><symbol value="PROCN000"/></f></fs></w> <w lemma="yo">yo<fs type="pos"><f name="P"><symbol value="PP1CSN00"/></f></fs></w> <w lemma="para">para<fs type="pos"><f name="S"><symbol value="SPS00"/></f></fs></w> <w lemma="mi"><choice><orig>mi</orig><reg>mí</reg></choice><fs type="pos"><f name="P"><symbol value="PP1CS000"/></f></fs></w> <w lemma="desear">deseo<fs type="pos"><f name="V"><symbol value="VMIP1S0"/></f></fs></w><pc change=". " type="F" subtype="Fp"/></seg> <s><w lemma="manuela" rend="manuela"><choice><orig>manu</orig><add hand="MS7" place="supralinear">e</add><lb break="no"/>la</orig><reg>Manuela</reg></choice><fs type="pos"><f name="N"><symbol value="NP00000"/></f></fs></w><pc change=". " type="F" subtype="Fc"/> <w lemma="este"><choice><orig>esta</orig><reg>ésta</reg></choice><fs type="pos"><f name="P"><symbol value="PD0FS000"/></f></fs></w> <w lemma="sólo"><choice><orig>solo</orig><reg>sólo</reg></choice><fs type="pos"><f name="R"><symbol value="RG"/></f></fs></w> <w lemma="servir"><choice><orig>sirbe</orig><reg>sirve</reg></choice><fs type="pos"><f name="V"><symbol value="VMIP3S0"/></f></fs></w> <w lemma="para">para<fs type="pos"><f name="S"><symbol value="SPS00"/></f></fs></w> <w rend="decirte">decirte<w part="I" lemma="decir">fs type="pos"><f name="V"><symbol value="VMN0000"/></f></fs></w> <w part="F" lemma="te">fs type="pos"><f name="P"><symbol value="PP2CS000"/></f></fs></w></w> <w lemma="donde"><choice><orig>Ande</orig><reg>ande</reg><sic>ande</sic></choice><fs type="pos"><f name="P"><symbol value="PT000000"/></f></fs></w> <w lemma="haber"><choice><orig>as</orig><reg>has</reg></choice><fs type="pos"><f name="V"><symbol value="VAIP2S0"/></f></fs></w> <w lemma="ir">ido<fs type="pos"><f name="V"><symbol value="VMP0000"/></f></fs></w> <w lemma="y">y<fs type="pos"><f name="C"><symbol value="CC"/></f></fs></w> <lb break="yes"/> <w lemma="cómo"><choice><orig>como</orig><reg>cómo</reg></choice><fs type="pos"><f name="P"><symbol value="PT000000"/></f></fs></w> <w lemma="te">te<fs type="pos"><f name="P"><symbol value="PP2CS000"/></f></fs></w> <w lemma="haber"><choice><orig>an</orig><reg>han</reg></choice><fs type="pos"><f name="V"><symbol value="VAIP3P0"/></f></fs></w> <w lemma="recibir">recibido<fs type="pos"><f name="V"><symbol value="VMP0000"/></f></fs></w><pc change=". " type="F" subtype="Fc"/> <w lemma="con">con<fs type="pos"><f name="S"><symbol value="SPS00"/></f></fs></w> <w lemma="buena">buena<fs type="pos"><f name="A"><symbol value="AQ0FS0"/></f></fs></w> <w lemma="cara">cara<fs type="pos"><f name="N"><symbol value="NCFS000"/></f></fs></w> <w lemma="o">o<fs type="pos"><f name="C"><symbol value="CC"/></f></fs></w> <w lemma="mala">mala<fs type="pos"><f name="A"><symbol value="AQ0FS0"/></f></fs></w> <w lemma="que">que</orig><reg>Que</reg></choice><fs type="pos"><f name="C"><symbol value="CS"/></f></fs></w> <w lemma="todo">todo<fs type="pos"><f name="P"><symbol value="PiOMS000"/></f></fs></w> <w lemma="lo">lo<fs type="pos"><f name="P"><symbol value="PP3CNA00"/></f></fs></w> <w lemma="quiero">quiero<fs type="pos"><f name="V"><symbol value="VMIP1S0"/></f></fs></w> <w lemma="saber">saber<fs type="pos"><f name="V"><symbol value="VMN0000"/></f></fs></w> <w lemma="porque">porque<fs type="pos"><f name="C"><symbol value="CS"/></f></fs></w> <w lemma="todavía"><choice><orig>todabía</orig><reg>todavía</reg></choice><fs type="pos"><f name="R"><symbol value="RG"/></f></fs></w> <w lemma="no">no<fs type="pos"><f name="R"><symbol value="RN"/></f></fs></w> <w lemma="me">me<fs type="pos"><f name="P"><symbol value="PP1CS000"/></f></fs></w> <lb break="yes"/> <w

```

¹⁹ For further information, see the attached table in 5.4.13.

```

lemma="haber"><choice><orig>e</orig><reg>he</reg></choice><fs type="pos"><f name="V"><symbol
value="VAIP1S0"/></f></fs></w> <w lemma="morir">muerito<fs type="pos"><f name="V"><symbol
value="VMP0000"/></f></fs></w> <w lemma="yo">yo<fs type="pos"><f name="P"><symbol
value="PP1CSN00"/></f></fs></w><pc change=". ." type="F" subtype="Fp"/></s>
<ss><w><choice><orig>Anque</orig><reg>Aunque</reg></choice><fs type="pos"><f name="C"><symbol
value="CC"/></f></fs><note type="linguisticNote" subtype="diphthong_and_hiatus"/></w> <w
lemma="estar"><choice><orig>estas</orig><reg>estás</reg></choice><fs type="pos"><f name="V"><symbol
value="VAIP2S0"/></f></fs></w> <w><choice><orig>trenita</orig><reg>treinta</reg></choice><fs type="pos"><f
name="Z"><symbol value="Z"/></f></fs><note type="linguisticNote" subtype="complex_syllable"/></w> <w
lemma="y">y<fs type="pos"><f name="C"><symbol value="CC"/></f></fs></w> <w lemma="2">dos<fs type="pos"><f
name="Z"><symbol value="Z"/></f></fs></w> <w lemma="legua">leguas<fs type="pos"><f name="N"><symbol
value="NCFP000"/></f></fs></w> <lb break="yes"/> <w
lemma="desviar"><choice><orig>desbiada</orig><reg>desviada</reg></choice><fs type="pos"><f name="V"><symbol
value="VMP00SF"/></f></fs></w> <w lemma="de">de<fs type="pos"><f name="S"><symbol
value="SPS00"/></f></fs></w> <w lemma="mi">mi</orig><reg>mí</reg></choice><fs type="pos"><f
name="P"><symbol value="PP1CSO00"/></f></fs></w>
```

5. Conversion to TEI-P5

5.1. General explanation

As it was previously explained in the Introduction, the XML encoding for *P.S. Post Scriptum* uses the model DALF (TEI-P4) as a starting point (internal model). This model, and, therefore, the corresponding DTD were customized in order to respond to *P.S. Post Scriptum* specific needs, meaning that they are only valid for internal use.

That internal model is available online, but it is also possible to download an up-to-date standardized model (external model). This last one is based on the model proposed by Red CHARTA, following the guidelines of TEI-P5. We offer below a detailed example of both models, internal and external, stressing the main differences between the two.

5.2. Internal model

Example from letter PS6235.

```
<TEI.2>
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>1650. Carta de José Antonio Díez para su hermano Manuel Díez, capellán.</title>
        <sponsor>CLUL, Centro de Linguística da Universidade de Lisboa</sponsor>
        <funder>ERC, European Research Council - 7FP/ERC Advanced Grant 2011 GA295562</funder>
        <respStmt><resp subcat="project"><name id="PS">PS, Post Scriptum</name></resp></respStmt>
        <respStmt><resp subcat="coordination"><name id="RM">Rita Marquilhas</name></resp></respStmt>
        <respStmt><resp subcat="transcription"><name id="GV">Gael Vaamonde </name></resp></respStmt>
        <respStmt><resp subcat="posannotation"><name/></resp></respStmt>
        <respStmt><resp subcat="parsing"><name/></resp></respStmt>
        <respStmt><resp subcat="modernizedEdition"><name>Gael Vaamonde</name></resp></respStmt>
        <respStmt><resp subcat="revision"><name/></resp></respStmt>
        <respStmt><resp subcat="translation"><name/></resp></respStmt>
        <respStmt><resp subcat="keywords"><name/></resp></respStmt>
        <respStmt><resp subcat="contextualization"><name>Gael Vaamonde</name></resp></respStmt>
      </titleStmt>
      <publicationStmt>
        <publisher>CLUL, Centro de Linguística da Universidade de Lisboa</publisher>
        <address>
          <addrLine>Av. Professor Gama Pinto, 2</addrLine>
          <addrLine>1649-003 Lisboa-Portugal</addrLine>
        </address>
        <pubPlace>Lisboa</pubPlace>
        <date value="2014">2014</date>
        <availability status="restricted"><p>Copyright 2014, CLUL</p></availability>
      </publicationStmt>
      <sourceDesc>
        <letDesc>
          <letIdentifier>
            <country>España</country>
            <settlement>Valladolid</settlement>
            <institution type="regionInstitute">Archivo de la Real Chancillería de Valladolid</institution>
            <repository reg="SC">Sala de lo criminal</repository>
          </letIdentifier>
        </letDesc>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
</TEI.2>
```

```

<collection>Pleitos criminales</collection>
<idno>Caja 766, Expediente 1</idno>
<altName><xref/></altName>
</letIdentifier>
<letHeading>
<author attested="yes" key="CDD.xml#JAD2" id="JAD2">José Antonio Díez</author>
<addressee attested="yes" key="CDD.xml#MD5">Manuel Díez</addressee>
<placeLet attested="yes">La Rioja, Calahorra</placeLet>
<placeDest attested="yes">Soria, El Burgo de Osma</placeDest>
<dateLet attested="yes">1807.06.27<date value="1807"/></dateLet>
</letHeading>
<physDesc>
<type>carta</type>
<support><p>un folio de papel doblado en cuarto, escrito por el recto del primer cuarto y por el recto del segundo cuarto.</p></support>
<extent>
<dimensions><height units="mm">201</height><width units="mm">191</width></dimensions>
</extent>
<layout><p>sin cambio de línea entre la fórmula de saludo y el inicio del texto.</p></layout>
<decoration>
<decoList>
<decoItem id="fig1"><decoDesc><p>cruz</p></decoDesc></decoItem>
</decoList>
</decoration>
<condition><p>/</p></condition>
</physDesc>
<envOcc occ="yes"/>
<letContents>
<locus>[13a]r y 14r</locus>
<class subcat="pragmatics">noticias</class>
<class subcat="type">familiar</class>
<class subcat="balancedSelection">1</class>
<class subcat="linguisticSource"/>
<class subcat="socioHistoricalSource"/>
<p>El autor escribe a su hermano para facilitarte la descripción de cierto sujeto.</p>
</letContents>
<additional>
<adminInfo>
<availability><p>Para hacer uso de las imágenes es necesaria la autorización del Archivo de la Real Chancillería de Valladolid.</p></availability>
</adminInfo>
<surrogates>
<p>
<bibl>facsímile digital guardado como fichero JPEG</bibl>
<bibl><xref subcat="PS6235_1.JPG"/></bibl>
<bibl><xref subcat="PS6235_2.JPG"/></bibl>
</p>
</surrogates>
<note subcat="context"><p>Proceso de 1807 sobre hurto dividido en dos piezas. La primera pieza recoge el pleito contra Manuel García sobre hurto de caballerías, apresión de ganzúa y llaves, y otros excesos. La segunda pieza, en la que se encuentra la carta aquí transcrita, recoge el pleito contra Manuel García sobre robo de una mula. </p>
<p>En 1807 apareció un hombre en Burgos sin papeles ni pasaporte, por lo que fue arrestado. El hombre dijo llamarse Francisco Gonzalo y ser labrador natural de Lerma (Burgos). Sin embargo, tras algunas averiguaciones y tras el reconocimiento de varios testigos se comprobó que en realidad se llamaba Manuel García, vecino de Medina de Pomar (Burgos), y que se había fugado de la cárcel de Miranda del Ebro. Además, se descubrió que era el mismo sujeto que en 1805 había robado una mula a José Antonio Díez, vecino de Calahorra (La Rioja), con el pretexto de ir a San Esteban de Gormaz (Soria) a llevar un poco de grana. La carta aquí transcrita fue aportada por Manuel Díez, capellán de la catedral de Burgos y hermano de José Antonio Díez, para demostrar que el reo era en efecto el ladrón de la mula de su hermano, pues en dicha carta se informaba de las señas del sujeto.</p></note>
<note subcat="translations" type="English">
<p subcat="title"></p>
<p subcat="summary"></p>
<p subcat="context"></p>
</note>

```

```

        </additional>
    </letDesc>
</sourceDesc>
</fileDesc>
<encodingDesc key="PS-GENERAL.xml#PS-ES">
    <projectDesc><p></projectDesc>
    <editorialDecl><p></editorialDecl>
</encodingDesc>
<profileDesc>
    <langUsage><language subcat="ES"/></langUsage>
</profileDesc>
<revisionDesc>
    <change when="2014-04">Gael Vaamonde: Codificación XML y contextualización</change>
    <change when="2014-10">Gael Vaamonde: Edición modernizada</change>
</revisionDesc>
</teiHeader>
<text id="PS6235">
    <envelope type="superscript">
        <envPart>
            <address type="receiver">
                <addrLine><deco decoRef="fig1"/></addrLine>
                <addrLine>A <abbr>Dn</abbr> Manuel Diez <abbr>contbro</abbr></addrLine>
                <addrLine>de <abbr>H</abbr> Por Soria</addrLine>
                <addrLine>el Burgo de Osma</addrLine>
            </address>
        </envPart>
    </envelope>
    <body><pb n="14r"/>
        <opener>
            <salute><deco decoRef="fig1"/></salute> <lb/>
            <letAddress>Calahorra</letAddress> <letDate>y Junio 27 de 1807</letDate>
        </opener>
        <p><salute>querido hermano</salute> qe||que el no aberme presentado <lb/> en esa a sido por estar con mucha
Familia y no estar <lb/> para acer gastos pues faltando a mi trabajo no <lb/> tengo qe||que comer tanbién conozco
qe||que no ago falta po<lb subcat="false"/>rqe||porque tu lo conoces lo mismo qe||que yo pues ablases barias <lb/>
beces co<del hand="JAD2">r</del>n el y aber tratado mientras estubo en esa <lb/> las señas d el son estas su
estatura como dos baras <lb/> poco mas mui rrojo y mui blanco mui gureso y mui <lb/> ancho de espaldas traia
sombbrero de copa alta re<lb subcat="false"/>decilla chaleco blanco de alcotonía calzones negros <lb/> medias de
algodon y apargatas balencianas y e labio <lb/> bajero partido por medio el nombre de el puedo jurar <lb/> qe||que me
dijo se llamaba Franco||Francisco el <gap reason="cancelled" extent="1 word"/> no estoí seguro <lb/> si me dijo se
llamababa Gonzalo o artiaga: espresio<lb subcat="false"/>nes de todos los de casa y manda a tu hermano
<salute>qe||que t||tu M||Mano B||Besa</salute></p>
        <closer>
            <signed>Josef Anto||Antonio Diez</signed> <lb/>
            <salute subcat="addressee">querido hermano</salute>
        </closer>
        <ps><p></p></ps>
    </body>
</text>
</TEI.2>

```

5.3. External model (TEI-P5)

Example taken from letter PS6235:

```
<?xml version="1.0" encoding="UTF-8"?>

<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:id="ES">

<teiHeader>

<fileDesc>

<titleStmt>

<title>1807. Carta de José Antonio Díez para su hermano Manuel Díez, capellán.</title>

<sponsor>CLUL, Centro de Linguística da Universidade de Lisboa</sponsor>

<funder>ERC, European Research Council - 7FP/ERC Advanced Grant 2011 GA295562</funder>

<respStmt><resp>Proyecto</resp><name xml:id="PS">PS, Post Scriptum</name></respStmt>

<respStmt><resp>Coordinación</resp><name xml:id="RM">Rita Marquilhas</name></respStmt>

<respStmt><resp>Transcripción</resp><name xml:id="GV">Gael Vaamonde</name></respStmt>

<respStmt><resp>Anotación POS</resp><name /></respStmt>

<respStmt><resp>Parsing</resp><name /></respStmt>

<respStmt><resp>Revisión</resp><name /></respStmt>

<respStmt><resp>Edición modernizada</resp><name>Gael Vaamonde</name></respStmt>

<respStmt><resp>Palabras-clave</resp><name /></respStmt>

<respStmt><resp>Contextualización</resp><name>Gael Vaamonde</name></respStmt>

</titleStmt>

<publicationStmt>

<publisher>CLUL, Centro de Linguística da Universidade de Lisboa</publisher>

<address>

<addrLine>Av. Professor Gama Pinto, 2</addrLine>

<addrLine>1649-003 Lisboa-Portugal</addrLine>

</address>

<pubPlace>Lisboa</pubPlace>

<date when="2014">2014</date>

<availability status="restricted"><p>Copyright 2014, CLUL</p></availability>

</publicationStmt>

<sourceDesc>

<msDesc>

<msIdentifier>

<country>España</country>

<settlement>Valladolid</settlement>
```

```
<institution key="CDD.xml#RegARCV">Archivo de la Real Chancillería de Valladolid</institution>

<repository xml:id="SC">Sala de lo criminal</repository>

<collection>Pleitos criminales</collection>

<idno>Caja 766, Expediente 1</idno>

</msIdentifier>

<msContents>

<summary>El autor escribe a su hermano para facilitarte la descripción de cierto sujeto.</summary>

<msItem class="original">

<p><locus>[13a]r y 14r</locus></p>

</msItem>

</msContents>

<physDesc>

<objectDesc form="letter">

<supportDesc>

<extent>un folio de papel doblado en cuarto, escrito por el recto del primer cuarto y por el recto del segundo cuarto.

<dimensions scope="all" unit="mm">

<height>201</height>

<width>191</width>

</dimensions>

</extent>

<condition><p></p></condition>

</supportDesc>

<layoutDesc>

<layout>sin cambio de línea entre la fórmula de saludo y el inicio del texto.</layout>

</layoutDesc>

</objectDesc>

<accMat>

<p>Sobrescrito:</p>

<p><quote xml:lang="ES">A Dn Manuel Diez contrbro<br/>de H Por Soria<br/>el Burgo de Osma</quote></p>

<p>Matasellos:</p>

<p><quote xml:lang="ES"/></p>

</accMat>

</physDesc>

<additional>

<adminInfo>

<availability><p>Para hacer uso de las imágenes es necesaria la autorización del Archivo de la Real Chancillería de Valladolid.</p></availability>
```

```

</adminInfo>

<surrogates>
<ref/>
<cit><bibl>facsímile digital guardado como fichero JPEG</bibl></cit>
</surrogates>
</additional>
</msDesc>
</sourceDesc>
</fileDesc>
<encodingDesc corresp="PS-GENERAL.xml#PS-ES">
<projectDesc><p/></projectDesc>
<editorialDecl><p/></editorialDecl>
<classDecl>
<taxonomy xml:id="Tip-PS">
<bibl>Tipología PS</bibl>
<category xml:id="pragmatics"><catDesc>noticias</catDesc></category>
<category xml:id="type"><catDesc>familiar</catDesc></category>
<category xml:id="linguisticSource"><catDesc/></category>
<category xml:id="socioHistoricalSource"><catDesc/></category>
<category xml:id="balancedSelection"><catDesc>1</catDesc></category>
</taxonomy>
</classDecl>
</encodingDesc>
<profileDesc>
<langUsage><language ident="ES"/></langUsage>
<textClass>
<catRef target="#pragmatics"/>
<catRef target="#type"/>
<catRef target="#linguisticSource"/>
<catRef target="#socioHistoricalSource"/>
<catRef target="#balancedSelection"/>
</textClass>
<creation/>
<correspDesc>
<correspAction type="sent">
<persName key="CDD.xml#JAD2" xml:id="JAD2" cert="high"><name>José Antonio Díez</name></persName>

```

```

<placeName confidence="1" type="origin">España, La Rioja, Calahorra
<location><geo>42.3073551 -1.9673133</geo></location>
</placeName>
<date confidence="1" when="1807-06-27" from="1807" to="1807" when-custom="1807"/>
</correspAction>
<correspAction type="received">
<persName key="CDD.xml#MD5" xml:id="MD5" cert="high"><name>Manuel Díez</name></persName>
<placeName confidence="1" type="destination">España, Soria, El Burgo de Osma</placeName>
<location><geo/></location>
</placeName>
</correspAction>
</correspDesc>

<settingDesc><p>Proceso de 1807 sobre hurto dividido en dos piezas. La primera pieza recoge el pleito contra Manuel García sobre hurto de caballerías, apresión de ganzúa y llaves, y otros excesos. La segunda pieza (en la que se encuentra la carta aquí transcrita) recoge el pleito contra Manuel García sobre robo de una mula.</p><p>En 1807 apareció un hombre en Burgos sin papeles ni pasaporte, por lo que fue arrestado. El hombre dijo llamarse Francisco Gonzalo y ser labrador natural de Lerma (Burgos). Sin embargo, tras algunas averiguaciones y tras el reconocimiento de varios testigos se comprobó que en realidad se llamaba Manuel García, vecino de Medina de Pomar (Burgos), y que se había fugado de la cárcel de Miranda del Ebro. Además, se descubrió que era el mismo sujeto que en 1805 había robado una mula a José Antonio Díez, vecino de Calahorra (La Rioja), con el pretexto de ir a San Esteban de Gormaz (Soria) a llevar un poco de grana. La carta aquí transcrita fue aportada por Manuel Díez, capellán de la catedral de Burgos y hermano de José Antonio Díez, para demostrar que el reo era en efecto el ladrón de la mula de su hermano, pues en dicha carta se facilitan las señas del sujeto.</p>
</settingDesc>
</profileDesc>
<revisionDesc>
<change when="2014-04">Gael Vaamonde: Codificación XML y contexto situacional</change>
<change when="2014-10">Gael Vaamonde: Edición modernizada</change>
<change xml:lang="EN">
<p sameAs="title"></p>
<p sameAs="summary"></p>
<p sameAs="settingDesc"></p>
</change>
</revisionDesc>
</teiHeader>
<text xml:id="PS6235">
<body><pb n="14r" facs="PS6235_2.JPG"/>
<opener>
<salute><seg type="sign">cruz</seg></salute> <lb/>
<placeName>Calahorra</placeName> y
<date> Junio 27 de 1807</date>
</opener>

```

```
<p><seg type="opener" subtype="salute">querido hermano</seg> [...]</p>
</body>
</text>
</TEI>
```

5.4. Main differences between the two models

5.4.1. Editing tasks

- Internal model. They are encoded as values (in English) of an attribute **@subcat** inside **<resp>**:

```
<respStmt><resp subcat="transcription"><name id="GV">Gael Vaamonde </name></resp></respStmt>
```

- External model. They are encoded as the content of an element **<resp>**:

```
<respStmt><resp>Transcripción</resp><name xml:id="GV">Gael Vaamonde</name></respStmt>
```

5.4.2. Communicative framework

- Internal model. They are encoded in **<fileDesc>** inside the element **<letHeading>**:

```
<fileDesc><sourceDesc><letDesc>
  <letHeading>
    <author attested="yes" key="CDD.xml#JAD2" id="JAD2">José Antonio Díez</author>
    <addressee attested="yes" key="CDD.xml#MD5">Manuel Díez</addressee>
    <placeLet attested="yes">La Rioja, Calahorra</placeLet>
    <placeDest attested="yes">Soria, El Burgo de Osma</placeDest>
    <dateLet attested="yes">1807.06.27<date value="1807"/></dateLet>
  </letHeading>
</letDesc></sourceDesc></fileDesc>
```

- External model. They are encoded in **<profileDesc>** inside the element **<correspDesc>**²⁰:

```

<profileDesc>
  <correspDesc>
    <correspAction type="sentBy">
      <persName attested="yes" corresp="CDD.xml#JAD2" xml:id="JAD2">José Antonio Díez</persName>
      <placeName attested="yes">La Rioja, Calahorra</placeName>
      <date attested="yes" when="1807.06.27"/>
    </correspAction>
    <correspAction type="deliveredTo">
      <persName attested="yes" corresp="CDD.xml#MD5">Manuel Díez</persName>
      <placeName attested="yes">Soria, El Burgo de Osma</placeName>
    </correspAction>
    <correspContext/>
  </correspDesc>
</profileDesc>

```

5.4.3. Summary of the letter

- Internal model. It is marked as the content of an element **<p>** inside **<letContents>**.

```

<fileDesc><sourceDesc><letDesc>
  <additional>
    <note subcat="context">
      <p>Redacción del contexto</p>
    </note>
    <note/>
  </additional>
</letDesc></sourceDesc></fileDesc>

```

²⁰ In the case of scribe or annotator intervention, their identification is to be declared, in the external model, inside **<profileDesc>** within the element **<creation>**:

```

<creation>
  <roleName xml:id="scribe"/>
  <roleName xml:id="annotator"/>
</creation>

```

- External model. It is marked as the content of the element **<summary>** inside **<msContents>**:

```
<fileDesc><sourceDesc><letDesc>
  <letContents>
    <p>Redacción del resumen.</p>
  </letContents>
```

5.4.4. Context of the letter

- Internal model. It is marked as the content of an element **<note>** inside **<additional>**.

```
<fileDesc><sourceDesc><letDesc>
  <additional>
    <note subcat="context">
      <p>Redacción del contexto</p>
    </note>
    <note/>
  </additional>
</letDesc></sourceDesc></fileDesc>
```

- External model. It is marked as the content of an element **<abstract>** inside **<profileDesc>**:

```
<profileDesc>
  <settingDesc><p>Redacción del contexto</p></settingDesc>
</profileDesc>
```

5.4.5. Translation of the letter

- Internal model. It is marked as the content of element **<note>** inside **<additional>**.

```
<fileDesc><sourceDesc><letDesc>
  <additional>
    <note />
    <note subcat="translations" type="English">
      <p subcat="title">Traducción del título</p>
      <p subcat="summary">Traducción del resumen</p>
      <p subcat="context">Traducción resumida del contexto</p>
      <p subcat="text">Traducción del texto</p> NOTA: Este elemento es excepcional.
    </note>
  </additional>
</letDesc></sourceDesc></fileDesc>
```

- External model. It is marked as the content of an element **<change>** inside **<revisionDesc>**.

```
<revisionDesc>
  <change xml:lang="EN">
    <p sameAs="title">Traducción del título</p>
    <p sameAs="summary">Traducción del resumen</p>
    <p sameAs="settingDesc">Traducción resumida del contexto</p>
    <p sameAs="text">Traducción del texto</p> NOTA: Este elemento es excepcional
  </change>
</revisionDesc>
```

5.4.6. Image description

- Internal model. It is marked with the element **<decoration>** inside **<physDesc>**. In the body of the text it is declared using an empty element **<deco/>**.

```
<physDesc>
  <decoration>
    <decoList>
      <decoItem id="fig1"><decoDesc><p>cruz</p></decoDesc></decoItem>
    </decoList>
  </decoration>
</physDesc>
```

```
[...] <deco decoRef="fig1"/> [...]
```

- External model. It is declared directly in the text, using the element **<seg type="sign">**:

```
[...] <seg type="sign">cruz<seg> [...]
```

5.4.7. Content clasification

- Internal model. The several classifications are encoded with the element **<class>** inside **<letContents>**:

```
<fileDesc><sourceDesc><letDesc>
  <letContents>
    <class subcat="pragmatics">noticias</class>
    <class subcat="type">familiar</class>
    <class subcat="balancedSelection">1</class>
    <class subcat="linguisticSource"/>
    <class subcat="socioHistoricalSource"/>
  </letContents>
</letDesc></sourceDesc></fileDesc>
```

- External model. The several classifications are encoded with the element **<catDesc>** inside **<encodingDesc>**. Each classification is related to an element **<catRef>** inside **<profileDesc>**:

```
<encodingDesc>
  <classDecl>
    <taxonomy xml:id="Tip-PS">
      <bibl>Tipología PS</bibl>
      <category xml:id="pragmatics"><catDesc>noticias</catDesc></category>
      <category xml:id="type"><catDesc>familiar</catDesc></category>
      <category xml:id="balancedSelection"><catDesc>yes</catDesc></category>
      <category xml:id="linguisticSource"><catDesc/></category>
      <category xml:id="socioHistoricalSource"><catDesc/></category>
    </taxonomy>
  </classDecl>
</encodingDesc>
<profileDesc>
  <textClass>
    <catRef target="#pragmatics"/>
```

```

<catRef target="#type"/>
<catRef target="#balancedSelection"/>
<catRef target="#linguisticSource"/>
<catRef target="#socioHistoricalSource"/>
</textClass>
</profileDesc>

```

5.4.8. Facsimile images

- Internal model. They are numbered as values of an attribute @**subcat** inside the element **<xref>**, in a list included inside **<surrogates>**:

```

<fileDesc><sourceDesc><letDesc>
  <additional>
    <surrogates>
      <p>
        <bibl>facsímile digital guardado como fichero JPEG</bibl>
        <bibl><xref subcat="PS6235_1.JPG"/></bibl>
        <bibl><xref subcat="PS6235_2.JPG"/></bibl>
      </p>
    </surrogates>
  </additional>
</letDesc><sourceDesc></fileDesc>

```

- External model. They are directly declared in the text, using the attribute @**facs** inside the element **<pb>**:

```

<text xml:id="PS6235">
  <body><pb n="14r" facs="PS6235_2.JPG"/>
    <opener>
      <salute><seg type="sign">cruz</seg></salute> <lb/>
      <placeName>Calahorra</placeName>
      <date>y Junio 27 de 1807</date>
    </opener>
    <p><seg type="opener" subtype="salute">querido hermano</seg> [...]</p>
  </body>
</text>

```

5.4.9. Envelope

- Internal model. It is transcribed with the element `<envelope>` inside `<text>`.

```
<text id="PS6235">
  <envelope type="superscript">
    <envPart>
      <address type="receiver">
        <addrLine><deco decoRef="fig1"/></addrLine>
        <addrLine>A <abbr>Dn</abbr> Manuel Diez <abbr>contbro</abbr></addrLine>
        <addrLine>de <abbr>H</abbr> Por Soria</addrLine>
        <addrLine>el Burgo de Osma</addrLine>
      </address>
    </envPart>
  </envelope>
</text>
```

- External model. It is transcribed in the header, using an element `<quote>` inside `<accMat>`:

```
<fileDesc><sourceDesc><msDesc>
<physDesc>
<accMat>
<p>Sobrescrito:</p>
<p><quote xml:lang="es"><seg type="sign">cruz</seg> <lb/> A <expan>Dn</expan> Manuel Diez
<expan>contbro</expan> <lb/> de <expan>H</expan> por Soria <lb/> el Burgo de Osma</quote></p>
</accMat>
</physDesc>
</msDesc></sourceDesc></fileDesc>
```

5.4.10. Postscript

- Internal model. It is transcribed inside the element `<ps>`.

Example from letter PS4131:

```
<ps>
  <p>A Antonio mi <lb/> vezino y su Mu-<lb subcat="false"/>ger muchas cosas. <lb/> y á Juanita la <lb/>
  vezina. y á Juan <lb/> Rexas el criado del <lb/> comand<ex>an</ex>te y su muger <lb/> Amigo Juan
  Manuel García</p>
</ps>
```

- External model. It is transcribed inside the element `<postscript>`.

```
<postscript>
  <p>A Antonio mi <lb break="yes"/> vezino y su Mu-<lb break="no"/>ger muchas cosas. <lb
  break="yes"/> y á Juanita la <lb break="yes"/> vezina. y á Juan <lb break="yes"/> Rexas el criado del
  <lb break="yes"/> <expan>comand<ex>an</ex>te</expan> y su muger <lb break="yes"/> Amigo Juan
  Manuel Garcia</p>
</postscript>
```

5.4.11. Line breaks

- Internal model. By default, line breaks are encoded with `<lb>`; however, line breaks that split a word in two are encoded with `<lb subcat="false"/>`.

las señas d el son estas su estatura como dos baras `<lb>` poco mas mui rrojo y mui blanco mui gureso y mui `<lb/>` ancho de espaldas traia sombrero de copa alta re`<lb subcat="false"/>`decilla chaleco blanco de alcotonia calzones negros

- External model. By default, line breaks are encoded with `<lb break="yes"/>`; however, line breaks that split a word in two are encoded with `<lb break="no"/>`.

las señas d el son estas su estatura como dos baras `<lb break="yes"/>` poco mas mui rrojo y mui blanco mui gureso y mui `<lb break="yes"/>` ancho de espaldas traia sombrero de copa alta re`<lb break="no"/>`decilla chaleco blanco de alcotonia calzones negros

5.4.12. Redundant or superfluous text: `<surplus>`

- The catchwords that can be found before page breaks are marked up with the element `<surplus>`:
porque yo `<surplus reason="catchword">deseo</surplus>` `<pb n="34v"/>` deseo que VM tenga buen suceso en sus empresas.
- Also, those superfluous words, or words chunks, that in the Post Scriptum internal model are conjectured as abandoned material, which the author forgot to cancel, are signalled with the same element `<surplus>`:
porque yo `<surplus reason="repeated">de</surplus>` deseo que VM tenga buen suceso en sus empresas.

5.4.13. Others

- The attribute `@id` changes to `@xml:id` whenever it appears in `<teiHeaders>`, but it is excluded if it appears in `<text>`.
- The elements `<letDesc>`, `<letIdentifier>` and `<letContents>` change to `<msDesc>`, `<msIdentifier>` and `<msContents>`.
- The elements `<letAddress>` and `<letDate>` change to `<placeName>` and `<date>`.

- The elements `<s @id @formula>` change into `<seg @type @subtype>`.
- The element `<lb line_rend="verticalText"/>` changes to `<lb rend="verticalText"/>`

5.4.14. Attachment text encoding

CORRESPONDENCIES BETWEEN TEXT ENCODING IN TEITOK AND IN TEI-P5

	TEITOK	TEI-P5
Numeration	<p id="p-1">	<p>
Abbreviations	<tok id="w-1" form="Snra" fform="Senhora">Snra</tok>	<w> <choice> <abbr>Snra</abbr> <expan>Senhora</expan> </choice> <w>
Word standardization (except cases of enclisis and contractions)	<tok id="w-10" nform="muito">muyto</tok>	<w> <choice> <orig>muyto</orig> <reg>muito</reg> </choice> <w>
Dialecta variant of a word (except cases of enclisis and contractions)	<tok id="w-174" nform="coisa" dform="cousa">Couza</tok>	<w> <choice> <orig>Couza</orig> <sic>cousa</sic> <reg>coisa</reg> </choice> <w>
Word lemma (except cases of enclisis and contractions)	<tok id="w-149" lemma="carne">carne</tok>	<w lemma="carne">carne</w>
POS annotation of word (except cases of enclisis and contractions)	<tok id="w-134" mfs="NCMS000" pos="N">almocreve</tok>	<w lemma="almocreve">almocreve <fs type="pos"> <f name="N"> <symbol value="NCMS000"/> </f> </fs> <w>
Full example with normalization (except cases of enclisis and contractions)	<tok id="w-144" form="samana" nform="semana" pos="N" mfs="NCFS000" lemma="semana">sama<lb id="e-18"/>na</tok>	<w rend="samana" lemma="semana"> <choice> <orig>sama<lb xml:id="lb-18" break="no"/>na</orig> <reg>semana</reg> </choice> <fs type="pos"> <f name="N"> <symbol value="NCFS000"/> </f> </fs> <w>

Full example of encoding with enclisis	<pre><tok id="w-3806" dform="assentou-se" nform="sentou-se">acentousse <dtok id="d-3806-1" pos="V" mfs="VMIS3S0" lemma="assentar" form="acentou" dform="assentou" nform="sentou"/> <dtok id="d-3806-2" pos="P" mfs="PP3CN000" form="sse" nform="se" lemma="se"/> </tok></pre>	<pre><w> <choice> <orig>acentouss</orig> <reg>sentou-se</reg> <sic>sentou-se</sic> </choice> <w part="I" lemma="sendar" rend="acentou" change="sentou"> <fs type="pos"> <f name="V"> <symbol value="VMIS3S0"/> </f> </fs> </w> <w part="F" lemma="se" rend="ss" sameAs="se"> <fs type="pos"> <f name="P"> <symbol value="PP3CN000"/> </f> </fs> <note type="dialectLemma" subtype="assentou"/> </w> </w></pre>
Full example of contraction encoding	<pre><tok id="w-144">na <dtok id="d-144-1" form="n" fform="em" pos="S" mfs="SPC00" lemma="em" /> <dtok id="d-144-2" form="a" pos="D" mfs="DA0FS0" lemma="o"/> </tok></pre>	<pre><w> <w part="I" rend="n" sameAs="em" lemma="em"> <fs type="pos"> <f name="S"> <symbol value="SPS00"/> </f> </fs> </w> <w part="F" rend="a" lemma="o"> <fs type="pos"> <f name="D"> <symbol value="DA0FS0"/> </f> </fs> </w> </w></pre>
Added punctuation	<pre><tok nform=". " id="w-3810" pos="F" mfs="Fp" lemma="."><ee/></tok></pre>	<pre><pc change=". " type="F" subtype="Fp"/></pre>
Edited punctuation	<pre><tok nform=". " id="w-381" pos="F" mfs="Fp" lemma=". ">,</pre>	<pre><pc change=". " type="F" subtype="Fp">,</pc></pre>
Deleted punctuation	<pre><tok id="w-31" nform="--",></tok></pre>	<pre><pc change="--",></pc></pre>
Line break (between words)	<pre><lb id="e-55"/></pre>	<pre><lb break="yes"/></pre>
Line break (inside words)	<pre><lb subcat="false" id="e-55"/></pre>	<pre><lb break="no"/></pre>
Line transcription for vertical text	<pre><lb line_rend="verticalText" id="e-55"/></pre>	<pre><lb break="yes" rend="verticalText"/></pre>
Formulaic parts outside <p></p>	<pre><salute/> <letAddress/> <letDate/> <signed/></pre>	<pre><salute/> <placeName/> <date/> <signed/></pre>
Formulaic parts inside <p></p> (concerning opener)	<pre><salute/> <letAddress/> <letDate/></pre>	<pre><seg type="opener" subtype="salute"/> <seg type="opener" subtype="placeName"/> <seg type="opener" subtype="date"/></pre>
Formulaic parts inside <p></p> (concerning closer)	<pre><salute/> <letAddress/> <letDate/> <signed/></pre>	<pre><seg type="closer" subtype="salute"/> <seg type="closer" subtype="placeName"/> <seg type="closer" subtype="date"/> <seg type="closer" subtype="signed"/></pre>
Salute outside <p></p> (concerning the addressee)	<pre><salute subcat="addressee">...</salute></pre>	<pre><salute <w>...</w> <w>...</w> ... <note type="addressee"/> </salute></pre>

Salute inside <p></p> (concerning the addressee)	<salute subcat="addressee"/>	<seg type="salute" subtype="addressee"/>
Linguistic notes	<tok ... itags="verb_paradigm"/>	<w>... <note type="linguisticNote" subtype="verb_paradigm"/> </w>

6. The biographic database

All information regarding the participants of the letters (authors and addressees) is encoded in a XML-TEI P4 database called CDD.xml (CARDS Demographic Database).

In this database, which inherits its name from the project CARDS, each entry is independent from the XML files of each letter. Nonetheless, all collected data can be cross-searched using the online search tool, available at the *P.S. Post Scriptum* website.

In the following section we present the XML conventions used to encode the participants' biographic information within the CDD.

6.1. The header

The header of the CDD.xml file is located inside the element `<teiHeader>`. Its content is invariable and short, since the only information mentioned is the title of the database, the languages used and the name of the creator of the XML framework:

```
<teiHeader type="text">
  <fileDesc>
    <titleStmt><title>CARDS demographic database</title></titleStmt>
    <publicationStmt><p/></publicationStmt>
    <sourceDesc><p>No source, born digital.</p></sourceDesc>
  </fileDesc>
  <profileDesc>
    <langUsage>
      <language id="PT"/>
      <language id="ES"/>
    </langUsage>
  </profileDesc>
  <revisionDesc>
    <change>
      <date>2007-10-09</date>
      <respStmt>
        <name>Ron Van den Branden</name>
        <resp>markup</resp>
      </respStmt>
      <item>Created the initial framework.</item>
    </change>
  </revisionDesc>
</teiHeader>
```

6.2. The participants list

The remaining content of the file is located inside the element `<text>`, whose internal structure can be seen below. The complete list of participants is included inside the element `<listPerson>`:

```

<text>
  <body>
    <div>
      <listPerson></listPerson>
    </div>
  </body>
</text>

```

Inside **<listPerson>**, there is the element **<person>**, which contains the biographic information of each participant. The internal structure of the element **<person>** can be seen below:

```

<person id="" role="" sex="" age="" lang="">
  <persName>
    <name></name>
  </persName>
  <affiliation/>
  <birth/>
  <death/>
  <education n="learnedLanguage"/>
  <event><desc/></event>
  <faith/>
  <floruit></floruit>
  <langKnowledge>
    <langKnown n="native"/>
  </langKnowledge>
  <nationality n="birthplace"/>
  <residence n="primary" key=""></residence>
  <residence n="secondary"/>
  <socecStatus key="">
    <state type="marriage"><p/></state>
    <trait><desc/></trait>
  </socecStatus>
</person>

```

In order to cover the data referring to every participant, a few general aspects were taken into consideration:

- Only toponyms and anthroponyms are written with initial capital letter. The rest of the words are always written in initial small letters.
- Never use full stop before closing an element.
- The different sequences of information included inside the same element are separated by semicolon.
- Conjectured information is always written inside square brackets.
- Each XML element can only be declared once (i.e. it is not possible to repeat elements inside **<person>**).
- The order of the elements inside **<person>** is always the same.

6.2.1. <person>

- a. The element **<person>**, which includes the remaining biographic elements of each participant is composed by five attributes: **@id**, **@role**, **@sex**, **@age** y **@lang**:

```
<person id="MBP1" role="author" sex="m" age="36" lang="PT">
```

- b. The attribute **@id** gives information on the code associated to each participant.
- c. The attribute **@role** gives information on the participant's role. It can only have one of the three possible values: "author", "addressee", "author/addressee".
- d. The attribute **@sex** gives information about the sex of the participant. It can only have one of the three possible values: "m" (masculine), "f" (feminine), "u" (unknown).
- e. The attribute **@age** gives information about the participants' age. It can only have a numeric value.
- f. The attribute **@lang** gives information about the language used by each participant (language in which the person wrote or received the letter). It can only have one of two possible values: "ES" (Spanish), "PT" (Portuguese).

6.2.2. <persName>

- a. This element includes the element **<name>** which contains the name of the participant.

```
<persName>
    <name>Fernando Esteban del Castillo y Valle</name>
</persName>
```

- b. The participant's name should be as complete as possible and it should be written with standardized spelling. Titles or forms of address are not used: sir, mister, count, duke, etc.
- c. If the participant has an *alias*, this is mentioned, following his name, as in the example²¹:

²¹ In the XML transcription in Oxygen, this information appears in the element **<title>** after the name and inside brackets. For example: **<title>1739. Carta de Manuela Rodríguez (La Riosela) para...</title>**. (cf. 1.1.1.1.m)

```
<persName>
    <name>Manuela Rodríguez, alias La Riosela</name>
</persName>
```

- d. If the participant used a fake name, but the real name is known, in this element only the real name is added. Nevertheless, in the element **<event><desc>** the fact that the participant used a fake name is mentioned.

```
<event><desc>usaba el pseudónimo de José Sáez Cantos</desc></event>
```

- e. In case of dictated letters, the biographic file in the CDD concerns the mental author, not the scribe.

6.2.3. **<affiliation>**

- a. This element gives information about kinship relations.

Ejemplo 1:
<affiliation>hermano de Francisco Pérez; nieto de Rafael Narciso</affiliation>

Ejemplo 2:
<affiliation>hijo de Pedro Pérez y de María Asensio</affiliation>

- b. The friendship relationships are not included here, but in the element **<event>**. The same applies to matrimonial relationships, which are included in the element **<state>**.

6.2.4. **<birth> and <death>**

- a. These elements account for the years of birth and death of the participant. It can only contain dates.

Example 1:
<birth>1800</birth>

Example 2:
<death>1655</death>

- b. If the exact date is unknown but it is possible to conjecture an interval, the interval is included.

```
<birth>[1800-1810]</birth>
```

- c. If the exact date is known (year, month and day), it is included as follows:

```
<death>25-03-1765</death>
```

6.2.5. <education>

- a. This element, which includes an attribute `@n="learnedLanguage"`, gives information about the foreign languages known by the participant.
- b. Each language is associated with a code: Latin (LT); French (FR); Spanish (ES); English (EN); Arabic (AR); Italian (IT); Greek (GR); German (AL). This element is always filled with the corresponding code.

`<education n="learnedLanguage>EN</education>`

- c. If the participant knows several languages, the corresponding codes are separated by semicolon.

`<education n="learnedLanguage>LT; EN; AL</education>`

- d. Any other type of information about the participant's education (e.g. "doesn't know how to read or write") is not included here, but in the element `<event>`.

6.2.6. <event>

- a. This element, which includes the element `<desc>`, gives information about important events in a participant's life.
- b. The text included inside the element `<event>` should be written using the past tense. Each different event is separated from the previous one by a semicolon.

`<event><desc>en 1825 fue apresado por intento de extorsión; en 1826 fue degradado para Cabo Verde por tres años</desc></event>`

- c. Friendship relationships are also included in this element (and not in `<affiliation>`, which is only used for kinship relationships).
- d. If the participant was prosecuted for an offense, that event is also registered:

`<event><desc>en 1658 fue acusado de un delito de bigamia</desc></event>`

- e. In this element it can be included information of the type "couldn't read nor write", if it applies (by default, it is assumed that all participants could read and write):

`<event><desc>no sabía leer ni escribir</desc></event>`

6.2.7. <faith>

This element contains information on the religious beliefs of the participant. The cases to be included are the following: Old Christian, New Christian, Protestant, Muslim, and Jewish.

```
<faith>cristiano nuevo</faith>
```

6.2.8. <floruit>

This element contains information about the years during which the participant wrote and/or received letters. It can only contain dates.

Ejemplo1:
<floruit>1825</floruit>

Ejemplo2:
<floruit>1725-1734</floruit>

6.2.9. <langKnowledge>

- a. This element, which includes an element **<langKnown>** and the attribute **@n="native"**, gives information on the participant's native language. This element must be filled with the same codes used to fill the element **<education>** (cf. 6.2.5.).
- b. The information about the native language is only filled when that fact is considered relevant, that is to say, when it is different from the information collected in the element **@lang=""**.

```
<langKnowledge>
  <langKnown n="native" >FR</langKnown>
</langKnowledge>
```

6.2.10. <nationality> and <residence>

- a. The element **<nationality>**, which includes the attribute **@n="birthplace"**, gives information on the participant's place of birth.
- b. The element **<residence>** gives information on the place where the participant lived. Two **<residence>** elements were considered. The first is the element **<residence subcat="primary">**, which corresponds to the participant's main residence and the second is the element **<residence subcat="secondary">**, which corresponds to a temporary residence, if there was any (no mention to prison's places should be registered here: cf. f) below).

```
<residence n="primary">Sevilla</residence>
<residence n="secondary">México, Ciudad de México</residence>
```

- c. There can only be one primary residence and one secondary residence. In case there were more than two residences, that should be mentioned in the element **<event>**.

```
<event><desc>tuvo como residencias Madrid, Toledo y París</desc></event>
```

- d. In order to make statistical treatment easier, all elements containing geographical information, such as **<nationality>** and **<residence>**, obey to the following specific rules. These rules are also applied to the content of the elements **<placeLet>** and **<placeDest>** (cf. 1.1.3.2.2.c).

- The first geographical reference is always a province/county (for Spanish data) or a judicial district (for Portuguese data).²²

Example 1:

ES
<nationality n="birthplace">Valladolid</nationality>

Example 2:

PT
<nationality n="birthplace">Moncorvo</nationality>

- If a more specific provenance is known, it is added on the right, separated by a comma.

ES Example 1:

<nationality n="birthplace">Valladolid, Cogeces del Monte</nationality>

PT Example 2:

<nationality n="birthplace">Moncorvo, Mirandela</nationality>

- If the participant was born or lived in another country, the name of the country will be added first, and only after a more specific provenance is added:

<nationality n="birthplace">Francia, París</nationality>

This last rule is applied to Spanish letters participants who were born or who lived in Portugal, as well as to Portuguese letters participants who were born or who lived in Spain. This also includes the countries which were, at a certain time, part of the Spanish and Portuguese empires. They are to be mentioned first:

Example 1:
ES

²² For this last case, we use the Carta das Memórias Paroquiais de 1758, created by the project *Atlas: Cartografia Histórica*, from the Instituto de História Contemporânea of the FCSH, UNL (<http://atlas.fcsh.ull.pt/cartoweb35/atlas.php>). Cf. see the corresponding table in the file Portugal_1758_Comarcas_Concelhos_Freguesias_lat_long.pdf.

```
<nationality n="birthplace">México</nationality>
```

Example 2:
PT

```
<nationality n="birthplace">Brasil, Bahia</nationality>
```

- As far as the islands are concerned, the districts are ignored:

Example 1:

```
<nationality n="birthplace">Açores, Ilha X, local</nationality>
```

Example 2:

```
<nationality n="birthplace">Madeira, local</nationality>
```

- In the case of places that, at the time, belonged to a country different from the current one, there will be no reference to the country (Ceuta, Gibraltar, Goa, Damão, Diu, Malaca, ...):

Example 1:

```
<nationality n="birthplace">Gibraltar</nationality>
```

Example 2:

```
<nationality n="birthplace">Goa</nationality>
```

There is also no reference to countries that were born during the timeframe covered by *P.S. Post Scriptum* (XVI to XIX centuries):
Belgium, every Popular Republics, Indonesia, ...)

Correct:

```
<nationality n="birthplace">Gante</nationality>
```

Incorrect:

```
<nationality n="birthplace">Bélgica, Gante</nationality>
```

- e. To the elements **<nationality>** and **<residence>** it is automatically added an additional attribute **@key** which contains the geographical coordinates of that specific place:

Example 1:

```
<nationality n="birthplace" key="38.722268, -9.139338">Lisboa</nationality>
```

Example 2:

```
<residence n="primary" key="43.009753, -7.556759">Lugo</nationality>
```

Example 3:

```
<residence n="primary" key="38.725229, -9.146107">Lisboa, Rua de Santa Marta, n.º 30</residence>
```

- f. If the secondary residence of a participant corresponds to a prison, it should not be included in the biographic file.

6.2.11. <socecStatus>

- a. This element refers to the socio-economic status of the participants. This information is divided in two parts. On the one hand, the element **<socecStatus>** gives information on the rank, the occupation and the profession of the participant. On the other hand, the attribute **@key** refers to the social classification of the participant, assigning him a category from a restrict group of options.
- b. The social classification of the participants relies on the type of court a person could appeal to during the Early Modern Age, according to his status. Nine possibilities were considered:

- Nobility

`<socecStatus key="nobility">conde de Aranda; presidente del Consejo de Castilla</socecStatus>`

- Ecclesiastical

`<socecStatus key="ecclesiastical">párroco de la iglesia Santa María la Mayor</socecStatus>`

- Member of the Inquisition

`<socecStatus key="inquisitorial">médico; familiar del Santo Oficio</socecStatus>`

- Military

`<socecStatus key="military">soldado</socecStatus>`

- Knightly Orders

`<socecStatus key="knightlyOrders">caballero de la Orden de Calatrava</socecStatus>`

- University

`<socecStatus key="university">estudiante; colegial del Colegio Mayor de San Ildefonso</socecStatus>`

- Ordinary

`<socecStatus key="ordinary">labrador</socecStatus>23`

- Slaves

`<socecStatus key="slave">esclavo</socecStatus>24`

²³ In the term "ordinary" are included the members of craft fraternities and other organizations, such as professional associations. Although this kind of associations ensured certain privileges during the *Ancien Régime*, their nature wasn't strictly juridical. For that reason, they aren't considered a different category.

²⁴ A freedman doesn't belong to the "Slaves" category, but to the "Ordinary" category, since after becoming a free man his juridical status changed.

- Unknown

`<socecStatus key="unknown"></socecStatus>25`

- If only the occupation of a family member is known, that information is registered.

Example 1:

`<socecStatus key="ordinary">mujer de zapatero</socecStatus>`

Example 2:

`<socecStatus key="ordinary">viuda de marinero</socecStatus>`

Example 3:

`<socecStatus key="ordinary">hija de labrador</socecStatus>`

- Positions and ranks are also registered in this element.

6.2.12. `<state>`

- This element, which includes the attribute `@type="marriage"` and the element `<p>`, gives information on the marital status of the participant:

Example 1:

`<state type="marriage"><p>casada con Gaspar de Cerqueira</p></state>`

Ejemplo 2:

`<state type="marriage"><p>viudo de Isabel Pérez</p></state>`

Ejemplo 3:

`<state type="marriage"><p>soltero</p></state>`

- When the marital status is known, but the name of the spouse is unknown, only the marital status is registered.

`<state type="marriage"><p>viuda</p></state>`

- In the case of ecclesiastical participants, no information is registered here.

6.2.13. `<trait>`

This element, which includes the element `<desc>`, contains the physical description of the participant.

`<trait><desc>cuerpo delgado, moreno, pelo de color castaño, chupado de cara, ojos pardos</desc></trait>`

6.3. Example of a biographic record

Example 1:

```
<person id="PPD1" role="author" sex="m" age="36" lang="ES">
    <persName><name>Pedro Pablo Díez</name></persName>
    <affiliation>hijo de Domingo Antonio Díez y Ana María Sánchez; padre de Paula Díez</affiliation>
    <birth>1673</birth>
    <death/>
    <education n="learnedLanguage"/>
    <event>
        <desc>estudió gramática en Alcalá de Henares; en 1711 fue condenado por la Inquisición por el delito de iluso e
            iludente</desc>
    </event>
    <faith>cristiano viejo</faith>
    <floruit>1706</floruit>
    <langKnowledge>
        <langKnown n="native"/>
    </langKnowledge>
    <nationality n="birthplace" key="40.526008, -2.823433">Guadalajara, Alhondiga</nationality>
    <residence n="primary" key="39.902922, -3.625574">Toledo, Yepes</residence>
    <residence n="secondary"/>
    <socecStatus key="ordinary">boticario</socecStatus>
    <state type="marriage"><p>primer matrimonio con Teresa Pérez Vázquez; casado desde 1701 con Petronila
        González</p></state>
    <trait>
        <desc>de buen cuerpo, algo calvo, pelo castaño oscuro poco poblado, hoyo en la barba</desc>
    </trait>
</person>
```

Example 2:

```
<person id="MF3" role="author" sex="f" age="" lang="PT">
    <persName><name>Maria da Fonseca</name></persName>
    <affiliation>filha de Pedro Dias e de Antónia Rodrigues; mãe de Mariana; irmã de Manuel Rodrigues</affiliation>
    <birth/>
    <death/>
    <education n="learnedLanguage"/>
    <event><desc>casou na freguesia de São José (Lisboa) a 9 de outubro de 1634; deste matrimónio teve uma filha, que  
faleceu ainda criança; tinha tenda de frutas da terra, pão e louça; fugiu do primeiro marido, o qual, depois, voltou a casar,  
razão pela qual ameaça acusá-lo de bigamia</desc></event>
    <faith/>
    <floruit>1661</floruit>
    <langKnowledge><langKnown n="native"/></langKnowledge>
    <nationality n="birthplace">Lisboa</nationality>
    <residence n="primary" key="38.722268, -9.139338">Lisboa, São José</residence>
    <residence n="secondary" key="-22.906538, -43.173523">Brasil, Rio de Janeiro</residence>
    <socceStatus key="ordinary">tendeira</socceStatus>
    <state type="marriage"><p>casada com Pascoal Coutinho</p></state>
    <trait><desc>alta de corpo, boa estatura, magra e alva; aparentava ter entre 45 e 50 anos</desc></trait>
</person>
```