



INSTITUTO DA LINGUA GALEGA

Gondomar. Corpus dixital de textos galegos
da Idade Moderna
Transcrición paleográfica. Criterios (versión 1)

Rosario Álvarez
Ernesto González Seoane

2017

INSTITUTO DA LINGUA GALEGA
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

**Gondomar. Corpus dixital de textos galegos
da Idade Moderna
Transcripción paleográfica. Criterios (versión 1)**

**Rosario Álvarez
Ernesto González Seoane**

Este traballo desenvolveuse no marco do proxecto *Corpus dixital de textos galegos da Idade Moderna. Catalogación, multiedición, glosario e estudo*, financiado cunha axuda do Ministerio de Economía y Competitividad (FFI2013-47589-P).

Cada documento, xa sexa impreso ou manuscrito, deberá constar dun elemento <TEI> que conterá como atributo o identificador do documento. Este deberá coincidir co campo que na ficha de catalogación se denomina "Referencia" e permitirá non só identificar o texto de xeito inequívoco, senón tamén servir de elemento de enlace do texto coas referencias bibliográficas correspondentes ou con outros recursos asociados, en particular coa súa reprodución facsimilar.

O identificador está formado por unha abreviatura da denominación do corpus (GOND) seguida inmediatamente dun número de orde, que se deberá completar con ceros á esquerda ata formar un número de tres cifras:

GOND002

Este identificador será empregado tamén como nome do ficheiro xml que conteña o texto e dos ficheiros jpg correspondentes ás reproducións facsimilares de cada unha das páxinas do texto. Estes últimos (un ficheiro por páxina) levarán unha extensión, separada do identificador, coa indicación da páxina de que se trate:

GOND002_1r.jpg

Cando dun mesmo texto existan varias versións, asignaráselles a todas elas un mesmo identificador seguido das letras "b", "c", "d", etc.:

GOND048

GOND048b

GOND048c

O elemento <TEI> conterá á súa vez dous elementos, correspondentes, respectivamente, á cabeceira (<teiHeader>) e ao texto propiamente dito (<text>):

```
<TEI xml:id="GOND002">
  <teiHeader>
    [...]
  </teiHeader>
  <text>
    [...]
  </text>
</TEI>
```

1. A cabeceira

A cabeceira do documento (<teiHeader>) deberá conter todas as informacións que permitan identificar o texto electrónico, o seu autor, as diferentes persoas e entidades que tiveron algunha responsabilidade ou interviñeron no proceso de edición, as características da edición, os datos relativos á composición, publicación, transmisión e conservación do texto orixinal, así como á súa caracterización lingüística e tipolóxica.

A cabeceira dos textos de Gondomar conterá os seguintes elementos:

- O elemento <fileDesc>, que contén as informacións relativas á identificación, catalogación e caracterización tanto do texto electrónico como do texto orixinal (autoría, proceso de edición, composición, transmisión, publicación, etc.).
- O elemento <encodingDesc> que documenta a relación existente entre o texto electrónico e o texto orixinal.
- O elemento <profileDesc> que informa das características tipolóxicas e lingüísticas do texto.

Así, a cabeceira dun documento de Gondomar quedaría conformada do seguinte xeito:

```
<teiHeader>
  <fileDesc>
<!-- ... -->
  </fileDesc>
  <encodingDesc>
<!-- ... -->
  </encodingDesc>
  <profileDesc>
<!-- ... -->
  </profileDesc>
</teiHeader>
```

1.1. Identificación e catalogación do texto

Os datos relativos á identificación e catalogación do texto documéntanse a través do elemento <fileDesc>, que conterà á súa vez os seguintes elementos:

- O elemento <titleStmt>, no que se rexistrarán os datos identificativos do texto electrónico.
- O elemento <publicationStms>, que recolle os datos que identifican a entidade responsable da edición do texto electrónico.
- O elemento <seriesStmt>, que identifica a serie ou colección en que se edita o texto electrónico.
- O elemento <sourceDesc>, que rexistra os datos identificativos e catalográficos do texto orixinal empregado como base da edición electrónica.

Polo tanto, esta sección queda estruturada do seguinte xeito:

```
<fileDesc>
  <titleStmt>
  <!-- ... -->
  </titleStmt>
  <publicationStmt>
  <!-- ... -->
  </publicationStmt>
  <seriesStmt>
  <!-- ... -->
  </seriesStmt>
  <sourceDesc>
  <!-- ... -->
  </sourceDesc>
</fileDesc>
```

a) Identificación do texto electrónico

Os datos básicos que permiten identificar o texto electrónico documéntanse a través do elemento <titleStmt> que, á súa vez, contén os seguintes elementos:

- O elemento <title>, no que se indicará o título do texto en versión normalizada (**de acordo cos criterios para a normalización de títulos**).

- O elemento <author>, no que se indicará o autor do texto, tamén en versión normalizada (**de acordo cos criterios de normalización**). Se o autor é descoñecido, indicaráse "Anónimo".
- O elemento <respStmt>, que permitirá identificar as distintas persoas que interviñeron no proceso de edición do texto electrónico e as tarefas desenvolvidas por cada unha delas. Este elemento conterá un ou varios elementos <resp> cun atributo @n, a través do cal se indicará a tarefa realizada ("principal", para os directores do proxecto; "funder", para a entidade financiadora do proxecto; "edition", para as persoas responsables da edición interpretativa; "transcription", para as persoas responsables da transcrición paleográfica; "annotation", para as persoas responsables da lematización e anotación..), e, no seu interior, un elemento <name>, no que se consignará o nome da persoa ou entidade responsable de cada tarefa ou función.

A modo de exemplo, esta sección debería quedar etiquetada do seguinte xeito:

```
<fileDesc>
  <titleStmt>
    <title>Romancillo da cea do cordeiro</title>
    <author>Anónimo</author>
    <respStmt>
      <resp n="principal">
        <name>Álvarez, Rosario </name>;
        <name>González Seoane, Ernesto </name>
      </resp>
      <resp n="funder">
        <name>Ministerio de Economía y Competitividad (referencia
          FFI2013-47589-P) </name>
      </resp>
      <resp n="edition">
        <name>Álvarez, Rosario </name>;
        <name>González Seoane, Ernesto </name>
      </resp>
      <resp n="transcription">
        <name>Beis, Sandra </name>
      </resp>
      <resp n="annotation">
        <name>González Seoane, Ernesto </name>
      </resp>
    </respStmt>
  </titleStmt>
</fileDesc>
```

b) Identificación da entidade responsable da edición do texto electrónico

O obxectivo desta sección é documentar, a través dun elemento <publicationStmt>, os datos editoriais correspondentes ao texto electrónico (entidade ou institución responsable da edición, lugar e data de publicación, dispoñibilidade do texto, etc.). Os elementos previstos nesta sección son os seguintes:

- O elemento <publisher>, que permite rexistrar o nome da institución ou entidade responsable da edición do texto electrónico.
- O elemento <pubPlace>, no que se recolle o lugar de publicación do texto electrónico.
- O elemento <date>, no que se indica o ano de edición do texto electrónico.
- O elemento <availability>, no que se informará sobre a dispoñibilidade e condicións de uso do texto electrónico (restricións, licenzas, etc.).

A modo de exemplo, esta sección debería quedar etiquetada do seguinte xeito:

```
<fileDesc>
  <publicationStmt>
    <publisher>Instituto da Lingua Galega</publisher>
    <pubPlace>Santiago de Compostela</pubPlace>
    <date>2015</date>
    <availability>
      <p>Distribuído baixo Creative Commons Attribution Non-
        Commercial 3.0 Unported License</p>
    </availability>
  </publicationStmt>
</fileDesc>
```

c) Identificación da serie ou colección en que se inscribe o texto electrónico

Esta información introdúcese a través dun elemento <seriesStmt>, dentro do cal a indicación da serie ou colección se ofrece mediante un elemento <title>:

```
<fileDesc>
  <seriesStmt>
    <title>Gondomar. Corpus dixital de textos galegos da Idade
    Moderna</title>
  </seriesStmt>
</fileDesc>
```

d) Identificación e catalogación do texto orixinal

Esta sección ten por obxecto identificar e catalogar, mediante un elemento <sourceDesc>, o texto empregado como fonte para a edición do texto electrónico.

Para a definición dos campos que van integrar esta sección debe terse en conta, por unha banda, que os textos que forman parte do corpus poden proceder tanto de fontes impresas como manuscritas. Por outra banda, en calquera dos dous casos os textos poden constituír un obxecto individualizado (impreso ou manuscrito) ou ben formar dun obxecto máis amplo.

b.1. Fontes impresas

Os campos necesarios para a descrición de fontes impresas corresponden aos elementos agrupados baixo a etiqueta <biblStruct>, un dos elementos posibles dentro da etiqueta <sourceDesc>.

Debe terse en conta que a maior parte dos textos do corpus procedentes de fontes impresas son pezas, fragmentos ou partes contidas en obras máis amplas. En consecuencia, é necesario distinguir, por unha banda, unha serie de campos referidos ao texto propiamente dito, que se agruparán baixo un elemento <analytic>, e, por outra, unha serie de campos que se refiren á obra ou volume en que está contido o texto, ou ben ao texto cando este constitúe unha obra independente, que se agruparán baixo un elemento <monogr>.

Desta maneira, no caso de fontes impresas esta sección queda estruturada polos seguintes elementos:


```
<fileDesc>
  <sourceDesc>
    <biblStruct>
      <analytic>
<!-- ... -->
      </analytic>
      <monogr>
<!-- ... -->
      </monogr>
    </biblStruct>
  </sourceDesc>
</fileDesc>
```

Dentro do elemento `<analytic>`, en que se describe o texto propiamente dito, poden inserirse os seguintes elementos:

- O elemento `<author>`, no que se indicará o autor do texto tal e como figura mencionado nel. De se tratar dunha obra asinada cunha indicación non transparente ("Otro de la Compañía", "Un gallego fiel", etc.) ou cun pseudónimo ("Marcos da Portela"), consignarase o nome ou denominación empregada. Se non hai ningunha indicación de autoría, consignarase "Anónimo".
- O elemento `<title>`, no que se indicará o título orixinal do texto na forma en que figura neste. Se a peza non ten título, crearase un a partir das súas primeiras palabras ou do seu primeiro verso. Ademais, indicarse que o texto forma parte dunha obra maior mediante un atributo `@level`.
- O elemento `<date>`, que se cubrirá só cando o texto leve (por exemplo, mediante unha indicación ao pé) unha data diferente da data de edición da obra que o contén.
- Un elemento `<textLang>`, no que indicará a lingua ou linguas en que está escrito o texto.

A información sobre a obra na que se inscribe o texto (ou o texto cando este constitúe unha obra independente) rexistrarse, como xa queda dito, mediante un elemento `<monogr>`, que conterá á súa vez os seguintes elementos:

- O elemento <author>, no que se indicará o nome do autor da obra tal e como figura nela. No caso de textos anónimos ou asinados con pseudónimo actuarase como se sinalou anteriormente. No caso de antoloxías ou obras de autor múltiplo en que non conste o nome do editor ou compilador, indicárase "Varios" dentro deste campo.
- O elemento <editor>, no que se indicará, no caso de antoloxías e obras de autor múltiplo, o nome da persoa que actúa como editor ou compilador. Cando este non conste, indicárase "Varios" dentro do elemento <author>.
- O elemento <title>, no que se consignará o título do volume ou da publicación periódica de que forma parte o texto. Ademais, mediante o atributo @level indicárase que se trata dunha obra máis ampla na que está incluído o texto.
- O elemento <imprint>, no que se indicarán os datos relativos á publicación do texto. Para iso, no seu interior inseríranse os elementos <pubPlace>, para indicar o lugar de publicación; <publisher>, para rexistrar o nome da editora ou imprenta, e <date>, para consignar o ano de publicación da obra.
- O elemento <extent>, no que se indicará o número de páxinas que compoñen a obra.
- O elemento <textLang>, no que se indicará a lingua ou linguas empregadas na obra.
- O elemento <biblScope>, a través do cal se indicará en que lugar da obra (volume, número, páxinas) está localizado o texto.

A modo de exemplo, esta sección da cabeceira quedaría estruturada do seguinte xeito:

```
<fileDesc>
  <sourceDesc>
    <biblStruct>
      <analytic>
        <author>Noguerol, Joseph</author>
        <title level="a">Versos gallegos à nuestra Señora de Reza,
        por avernos dado un Príncipe en nuestra Reyna, y Señora
        María Gabriela, por un Autor no conocido, ni dado à luz por
        mano de D. Joseph Noguerol su mismissimo primo
        hermano</title>
        <textLang>Texto en galego, título en español</textLang>
      </analytic>
      <monogr>
        <editor>Cossío y Bustamante, Diego de</editor>
        <title level="m">El clarin de la fama y cítara de Apolo con
        métricos rasgos a las reales fiestas que en el felicísimo
        nacimiento de el Príncipe N. Señor D. Luis Jacobo Primero el
        deseado executó la esclarecida nobilísima, y muy Leal
        Ciudad de Orense, y oy consagra a la augusta sombra de la
        Rey N. Señora D. María Luisa Gabriela Enmanuel de
        Saboya</title>
        <imprint>
          <pubPlace>Santiago</pubPlace>
          <publisher>Imprenta de Antonio de Aldemunde</publisher>
          <date>1708</date>
        </imprint>
        <extent>15 f. + 215p.</extent>
        <textLang>Textos en español, latín e galego</textLang>
        <biblScope>pp. 179-180</biblScope>
      </monogr>
    </biblStruct>
  </sourceDesc>
</fileDesc>
```

b.2. Fontes manuscritas

Os elementos necesarios para a descrición de fontes manuscritas corresponden ás informacións agrupadas baixo o elemento <msDesc>, un dos elementos posibles dentro de <sourceDesc>.

O elemento <msDesc> pode, á súa vez, conter no seu interior os seguintes elementos:

- O elemento <msIdentifier>, que contén todas as informacións necesarias para identificar e localizar o manuscrito correspondente ao texto.
- O elemento <msContents>, que recolle a descrición da parte do manuscrito que contén o texto.

- O elemento <physDesc>, que permite realizar unha descrición completa das características físicas do manuscrito.
- O elemento <history>, no que se recolle a información sobre a historia do manuscrito.

Desta maneira, esta sección da cabeceira quedaría estruturada do seguinte xeito:

```
<fileDesc>
  <sourceDesc>
    <msDesc>
      <msIdentifier>
<!-- ... -->
      </msIdentifier>
      <msContents>
<!-- ... -->
      </msContents>
      <physDesc>
<!-- ... -->
      </physDesc>
      <history>
<!-- ... -->
      </history>
    </msDesc>
  </sourceDesc>
</fileDesc>
```

No interior do elemento <msIdentifier> intégranse unha serie de elementos que teñen por obxecto identificar e localizar o manuscrito correspondente ao texto (ou en que figura o texto):

- O elemento <institution>, no que se recolle o nome da institución en que se encontra depositado o manuscrito.
- O elemento <repository>, no que se recolle o nome da biblioteca ou arquivo no que se conserva o manuscrito.
- O elemento <settlement>, no que se rexistra o nome da cidade en que se encontra a dita institución.
- O elemento <country>, no que se rexistra o nome do país correspondente.
- O elemento <collection>, no que se indica o nome da colección de que forma parte o manuscrito.

- O elemento <idno>, no que se indica a signatura do manuscrito.
- O elemento <altIdentifier>, que pode conter no seu interior os elementos sinalados anteriormente, e que permite documentar o paso do manuscrito por outras institucións, así como signaturas que tivo anteriormente dentro da mesma institución.
- O elemento <msName>, no que se consignará o nome con que é coñecido o manuscrito que contén o texto.

A modo de exemplo, esta sección debería quedar estruturada do seguinte xeito:

```
<msDesc>
  <msIdentifier>
    <country>España</country>
    <settlement>Madrid</settlement>
    <institution>Real Academia de la Historia</institution>
    <repository>Biblioteca de la Real Academia de la
    Historia</repository>
    <collection>Papeles de José Cornide</collection>
    <idno>M-RAH 9/3894</idno>
    <altIdentifier>
      <country>Estados Unidos</country>
      <settlement>New York</settlement>
      <repository>Pierpont Morgan Library</repository>
      <idno>ms. 5840</idno>
    </altIdentifier>
    <msName>Papeles varios mss. : de diversas procedencias y
    asuntos, algunos de don José de Cornide, quien al parecer los
    reunió</msName>
  </msIdentifier>
</msDesc>
```

No interior do elemento <msContents> intégranse unha serie de elementos que teñen por obxecto describir o contido do manuscrito e, en particular, a parte do manuscrito que contén o texto (tanto se manuscrito e texto coinciden coma se non):

- O elemento <textLang>, no que se indica a lingua ou linguas empregadas no conxunto do manuscrito.
- O elemento <msItem>, que permite rexistrar as informacións relativas á parte do manuscrito que contén o texto. Para iso, no seu interior conterà os seguintes elementos:

- <author>, no que se consignará o nome do autor do texto na súa forma orixinal, non normalizada. Se o texto é anónimo, indicárase "Anónimo"; se está asinado baixo pseudónimo ou cunha atribución non transparente, consignárase esta.
- <title>, no que se consignará o título do texto na súa forma orixinal. Se o texto carece de título créase un a partir das súas primeiras palabras ou primeiro verso.
- <respStmt>, no que se consignarán outras mencións de responsabilidade. Con este fin, emprégase o atributo @n para indicar a función realizada ("edition", "transcription", "translation", etc.) e o elemento <name> para documentar o nome da persoa responsable.
- <locus>, no que se indicará a localización do texto dentro do manuscrito.
- <textLang>, no que se indicará a lingua ou linguas empregadas no texto.

A modo de exemplo, esta sección da cabeceira quedaría estruturada do seguinte xeito:

```
<msDesc>
  <msContents>
    <textLang>Textos en español, latín e galego</textLang>
    <msItem>
      <title>Señales que han de preceder el fin del mundo</title>
      <author>Sánchez, Antonio</author>
      <respStmt n="transcription">
        <name>Aguirre, Diego</name>
      </respStmt>
      <locus>fol. 14r-18v</locus>
      <textLang>Texto en galego, título en español</textLang>
    </msItem>
  </msContents>
</msDesc>
```

No interior do elemento <physDesc> intégranse unha serie de elementos que teñen por obxecto describir as características físicas do manuscrito:

- O elemento <objectDesc>, que contén á súa vez unha serie de elementos destinados a describir as características do manuscrito:
 - <supportDesc>, no que se recollen as informacións relativas aos aspectos materiais do manuscrito, da súa constitución, do seu estado de conservación, etc.
 - <support>, para indicar o tipo de soporte (papel, pergamiño...)
 - <extent>, para indicar o número de páxinas
 - <dimensions> (cun atributo @unit, no que se indicará a unidade de medida e uns elementos <height> e <width>), para indicar as medidas do manuscrito
 - <collation>, para indicar como está constituído fisicamente o manuscrito
 - <foliation>, para indicar o sistema empregado para a numeración das páxinas do manuscrito.
 - <condition>, para dar conta do estado de conservación do manuscrito
 - <layoutDesc>, no que se ofrecen detalles sobre a distribución do texto en páxina.
- O elemento <handDesc>, no que dará conta das mans que interviñeron no manuscrito.
- O elemento <bindingDesc>, no que se indicará o tipo de encadernación do manuscrito

A modo de exemplo, esta sección quedaría estruturada do seguinte xeito:

```
<msDesc>
  <physDesc>
    <objectDesc>
      <supportDesc>
        <support>Papel</support>
        <extent>248 páxinas</extent>
        <dimensions unit="cm">
          <height>22</height>
          <width>15</width>
        </dimensions>
        <collation>
          <p>O manuscrito presenta catro partes diferenciadas; a primeira, segunda e cuarta constan de 45 páxinas cada unha, mentres que a terceira, que parece unha interpolación posterior, está formada por 110 páxinas</p>
        </collation>
        <foliation>
          <p> Páxinas numeradas con arábigos correlativos, do 1 ao 247 no ángulo superior externo; a numeración é orixinal, concordante co índice</p>
        </foliation>
        <condition>
          <p>Bo estado de conservación; faltan as pp. 29-30, que foron arrincadas</p>
        </condition>
      </supportDesc>
      <layoutDesc>
        <p>Texto a dúas columnas, de entre 24 e 28 liñas</p>
      </layoutDesc>
    </objectDesc>
    <handDesc>
      <p>Letra de diferentes mans. Existe unha man principal, responsable da maior parte das páxinas do volume. Ademais, numera as páxinas e elabora o índice. As restantes mans interveñen na copia de poemas nos espazos deixados en branco pola man principal e introducen glosas e notas marxinais</p>
    </handDesc>
    <bindingDesc>
      <p>Encadernación en pergameo</p>
    </bindingDesc>
  </physDesc>
</msDesc>
```

Finalmente, no interior do elemento <history> intégranse unha serie de elementos que teñen por obxecto estruturar a información sobre a historia do manuscrito:

- O elemento <origin>, no que se indicará, se é posible, cando, onde e en que circunstancias foi redactado o manuscrito

- O elemento <provenance>, no que se documentará, se é posible, o percorrido do manuscrito desde a súa redacción ata a súa adquisición polo seu propietario actual
- O elemento <acquisition>, no que se indicará, se é posible, a fórmula ou procedemento polo que o manuscrito chegou ás mans do propietario actual.

Así, esta sección da cabeceira quedaría estruturada do seguinte xeito:

```
<msDesc>
  <history>
    <origin>
      <p>Escrito no mosteiro de Lézrez na segunda metade do século
        XVI</p>
    </origin>
    <provenance>
      <p>En data non determinada pasa a formar parte da biblioteca
        particular da familia Osorio</p>
    </provenance>
    <acquisition>
      <p>Doado en 2014 ao Museo de Pontevedra pola familia
        Osorio</p>
    </acquisition>
  </history>
</msDesc>
```

1.3. Relación entre texto electrónico e texto orixinal

Esta información consignarase a través dun elemento <encodingDesc>:

```
<teiHeader>
  <encodingDesc>
    <p>Transcripción conservadora do texto orixinal. Mantivéronse a
      grafía e acentuación orixinais. Regularizáronse os signos de
      puntuación.</p>
  </encodingDesc>
</teiHeader>
```

1.4. Clasificación lingüística e tipolóxica do texto

Para a clasificación lingüística e tipolóxica do texto empregase un elemento <profileDesc>, que á súa vez conterá no seu interior os seguintes elementos:

- O elemento <creation>, que conterá no seu interior un elemento <date> no que se indicará o ano no que se quere datar o texto. Este elemento será empregado logo como parámetro nas buscas que se realicen sobre o corpus. É necesario que neste campo se consigne un ano (sen ningunha indicación adicional), mesmo naqueles casos en que non sexa posible establecer con precisión e seguridade a súa data de composición.
- O elemento <textClass>, que conterá no seu interior un elemento <keywords>. No seu interior consignaranse, a través do elemento <term>, as características tipolóxicas do texto de acordo cos seguintes parámetros de clasificación¹:
 - @subcat="source type"
 - Impreso
 - Manuscrito
 - @subcat="text type"
 - Textos literarios
 - Poesía culta
 - Poesía popular
 - Poesía cancioneril
 - Vilancico
 - Texto dramático ou dialogado
 - Prosa
 - Textos non literarios
 - Prosa documental
 - Correspondencia privada
 - Texto historiográfico
 - Refrás e proverbios
 - Glosa ou anotación
 - @subcat="circulation"
 - Textos privados

¹ Cf. Rosario Álvarez/Ernesto González Seoane (2016): «Iluminar los *Séculos Escuros*: Gondomar, un corpus para el estudio del gallego en la Edad Moderna», in: Johannes Kabatek, ed. (no prelo).

- Textos semipúblicos
- Textos públicos
- @subcat="autonomy"
 - Textos autónomos
 - Textos compilados
 - Textos non autónomos
- O elemento <langUsage>, cun elemento <language> no seu interior, que permite identificar a lingua, variedade ou rexistro en que está escrito o texto. Sen prexuízo de que sexa necesario ampliar esta clasificación, prevense a priori as seguintes variedades:
 - Galego
 - Hibridación galego-castelán
 - Hibridación galego-portugués (base portuguesa)
 - Galego arcaizante
 - Galego paródico
 - Hibridación galego-portugués (base galega)

A modo de exemplo, esta sección da cabeceira quedaría estruturada do seguinte xeito:

```
<teiHeader>
  <profileDesc>
    <creation>
      <date>1697</date>
    </creation>
    <textClass>
      <keywords>
        <term subcat="source type">Impreso</term>
        <term subcat="text type">Textos literarios/Poesía
        culta</term>
        <term subcat="circulation">Textos públicos</term>
        <term subcat="autonomy">Textos compilados</term>
      </keywords>
    </textClass>
    <langUsage>
      <language>Galego</language>
    </langUsage>
  </profileDesc>
</teiHeader>
```

2. O texto

2.1. Elementos estruturais básicos

O texto propiamente dito quedará demarcado, como se indicou, por un elemento <text>, que conterá un elemento obrigatorio <body>. Este poderá ir precedido dun elemento opcional <front> e seguido dun elemento <back>, tamén opcional:

- O elemento <body> contén o texto propiamente dito.
- O elemento <front> empregárase para demarcar mencións que preceden o texto e que non forman parte del, como por exemplo, un título ou a indicación do tipo de estrofa de que se trata, o nome do autor, unha dedicatoria, etc. Nos textos deste período é frecuente que estas mencións estean escritas en castelán. Esta circunstancia indicárase mediante o elemento <foreign>, de acordo co que se establece en ***.
- O elemento <back> empregárase para demarcar mencións de características análogas ás mencionadas para <front> que aparecen a continuación do texto propiamente dito (por exemplo, nome do autor, lugar de redacción do texto, data...).

Deste xeito, a estrutura primaria dun texto quedaría conformada do seguinte xeito:

```
<text>
  <front>
    [...]
  </front>
  <body>
    [...]
  </body>
  <back>
    [...]
  </back>
</text>
```

O elemento <front> pode conter unha ou varias seccións, que se marcarán empregando o elemento <div>. Á súa vez, <div> pode conter un ou varios dos seguintes elementos (indícase simplemente unha relación provisoria):

- Un elemento <head> para a mención do título e autor do texto.
- Un elemento <epigraph> para as citas introdutorias, dedicatorias, etc.
- Un elemento <set> para a descrición da escena ou para a localización espacial ou temporal da acción (no caso das pezas teatrais ou parateatrais).
- Un elemento <castList> para introducir a relación de personaxes (e eventualmente o reparto de actores) (no caso de pezas teatrais ou parateatrais. Este elemento pode incluír, á súa vez, un elemento <head> para introducir a fórmula empregada para encabezar esta sección no texto orixinal e un elemento <castItem> para cada personaxe.

```
<front>
  <div>
    <head>
      <foreign xml:lang="es">SONETO GALLEGO <lb/>de el
        Licenciado Pedro Bazquez <lb/>de Neyra</foreign>
    </head>
  </div>
  <div>
    <epigraph>
      <p><lb/><foreign xml:lang="la">RESPICE
        FINEM</foreign></p>
    </epigraph>
  </div>
</front>
```

```
<front>
  <div>
    <head>
      <foreign xml:lang="es">Entremes famoso en que Contiene
        la <lb/> Contienda que tubieron los labradores dela
        <lb/>fra de Caldelas Con los portogueses sobre la pesca
        <lb/>del Rio miño. Año de 167= Compuesto por el <lb/>
        licenciado gabriel feyxò de arauso año <lb/> de
        1671</foreign>
    </head>
  </div>
  <div>
    <castList>
      <head><foreign
        xml:lang="es">Personas</foreign><lb/></head>
      <castItem><foreign xml:lang="es">Un
        portogues</foreign></castItem> = <castItem><foreign
        xml:lang="es">un labrador <lb/> llamado
        Roleiro</foreign></castItem> = <castItem><foreign
        xml:lang="es">Una portoguesa</foreign></castItem>
    </castList>
  </div>
</front>
```

O elemento <body> contén un ou varios elementos <div>, que corresponden ás divisións ou seccións que pode presentar o corpo do texto. Empregarase este elemento para diferenciar, por exemplo, partes do texto separadas por epígrafes, espazos en branco, sinais divisorios, etc. No caso de que o texto non presente divisións empregárase unicamente un elemento <div>.

Nos textos en prosa e verso, cada elemento <div> pode conter un ou varios dos seguintes elementos:

- <p> para parágrafos en prosa.
- <lg> para grupos de versos que funcionan como unha unidade formal (por exemplo, estrofas, retrosos, etc.). Os distintos versos contidos dentro dun elemento <lg> estarán demarcados polo elemento <l>.

Nos textos teatrais ou parateatrais, as diferentes unidades estruturais marcaranse mediante cadanseu elemento <div>, que, de ser o caso, deberá ir acompañado dun atributo @type no que se indicará o tipo de unidade de que se trate ("act", "scene"...), e dun atributo @n no que se indicará o número que corresponda. Cando o texto presenta varios niveis estruturais xerarquizados (por exemplo, "Acto I. Escena 2ª"), entregarase <div1> para as unidades do primeiro nivel e <div2> para as unidades de segundo nivel. Se o texto non presenta unidades estruturais, empregárase un único elemento <div>.

Nos textos dramáticos e, en xeral, nos textos dialogados, cada unha das intervencións deberá ir inserida dentro dun elemento <sp>, acompañado dun atributo @who, no que se especificará o nome ou a identidade do personaxe ou dos personaxes que falan. Esta identificación realizarase a través dunha forma normalizada e, por tanto, non necesariamente coincidente coa que figura no texto orixinal introducindo cada parlamento. Este elemento <sp> poderá, á súa vez, conter os seguintes elementos:

- <speaker> para a indicación do nome ou mención tal e como figura no texto orixinal.
- <p> para os parlamentos en prosa.

- `<l>` para os parlamentos en verso.
- `<lg>` para poemas, cancións ou fragmentos en verso inseridos no texto e que presentan entidade propia. Este elemento conterá no seu interior un elemento `<l>` para cada verso.

Nos textos teatrais e parateatrais, as indicacións escénicas (decorados, accións que realizan os personaxes, entradas e saídas de escena, etc.) márcanse mediante o elemento `<stage>`.

Deste xeito, o corpo dun texto en verso podería quedar conformado do seguinte xeito:

```
<body>
  <div>
    <lg>
      <l>Morte cruel esa tredora maña</l>
      <l>de roubar de non cato á humana vida</l>
      <l>conque ollos á podeche ver comprida</l>
      <l>na santa REINA que oge perde España?</l>
    </lg>
    <lg>
      [...]
    </lg>
    [...]
  </div>
</body>
```

Exemplo dun texto en prosa.

O corpo dun texto dramático quedaría conformado do seguinte xeito:

```
<body>
  <div1 type="act" n="1">
    <div2 type="scene" n="2">
      <stage>Entra correndo Perucho.</stage><lb/>
      <sp who="Perucho">
        <speaker>Per.</speaker>
        <p>Achégate acó, Mingote.</p><lb/>
      </sp>
      <sp who="Mingote">
        <speaker>Ming.</speaker>
        <p>E logo, que queres, ho?</p><lb/>
      </sp>
    </div2>
  </div1>
</body>
```

O elemento <back> pode, á súa vez, conter unha varias seccións, que se marcarán mediante o elemento <div>. Este elemento pode incluír no seu interior outros elementos, como por exemplo:

- <trailer> para indicacións de final de texto, colofóns, etc.

```
<back>
  <div>
    <trailer>
      <p>"<foreign xml:lang="la">finis Laus Deo</foreign></p>
    </trailer>
  </div>
</back>
```

2.2. Saltos de páxina, columna e liña

a) Saltos de páxina

O comezo dunha nova páxina ou folio márcase cun elemento <pb/>. Este elemento vai acompañado de dous atributos:

- Un atributo @n, que indica o número de páxina ou folio

```
<pb n="1r"/>
<pb n="235"/>
```

- Un atributo @facs, que permite vincular as páxinas do texto cos arquivos que conteñen as reproducións facsimilares correspondentes. Este atributo pódese xerar de maneira automática a través de TEITOK, polo que non hai necesidade de introduci-lo manualmente:

```
<pb n="1r" facs="GOND002_1r.jpg"/>
<pb n="235" facs="GOND045_235.jpg"/>
```

Independentemente do número de páxinas de que conste un texto, inserirase sempre un elemento <pb/> para marcar o seu comezo.

Cando o salto de páxina cadra no interior dunha palabra, o elemento <pb/> insírese no seu interior (e elimínase o guión de final de liña):

```
pala<pb n="235" facs="GONDO45_235.jpg"/>bra
```

b) Saltos de columna

Os saltos de columna márcanse inserindo no lugar correspondente un elemento <cb/>, seguido dun atributo @n que indica o número de columna. Para este fin empregaranse as letras "a" e "b" (e "c"... no caso improbable de que sexan máis de dúas).

```
<cb n="a"/>
```

```
<cb n="b"/>
```

c) Saltos de liña

Nos textos en prosa os saltos de liña márcanse inserindo no lugar correspondente un elemento <lb/>, seguido dun atributo @n que indica o número de liña.

```
Llegò fuera de el termino de el Certamen, y por <lb n="2"/>esso  
no se premiò en el lugar que merecia.<lb n="3"/>
```

Para a indicación das liñas nos textos en verso empregárase, segundo queda dito, un elemento <l>.

Nos casos en que o salto de liña cadre no interior dunha palabra, o elemento insírese no lugar correspondente (con eliminación do guión de final de liña):

```
pala<lb n="2"/>bra
```

2.3. O tratamento do texto. Instrucións para a transcrición paleográfica

A transcrición paleográfica ten por obxecto ofrecer unha reprodución fiel da grafía do texto orixinal. Por esta razón, respectaranse trazos coma os seguintes:

- Grafías
- Acentuación
- Puntuación
- Unión e separación de palabras
- Abreviaturas

Ao tempo, nesta transcrición sinalaranse lagoas, riscaduras, emendas, dificultades de lectura, etc. que se aprecien no texto orixinal.

a) Riscaduras e supresións existentes no texto orixinal

Os fragmentos que no texto orixinal aparezan riscados mediante un trazo ou un borrancho marcaranse mediante un elemento seguido dun atributo @type, no que se indicará "overstrike".

No caso de que o texto riscado sexa lexible, esta circunstancia sinalarase conforme se indica a seguir:

<del type="overstrike">ti non me has ti de aloumiñar

Se, pola contra, o texto riscado é ilexible, empregarase un elemento baleiro con indicación, se é posible, do número de caracteres riscados (mediante os atributos @quantity e @unit):

<del type="overstrike" quantity="4" unit="character"/>

Se o procedemento empregado para o riscado fose outro diferente, deberase indicar empregando tamén o atributo @type ("cancelado", "raspado"...) e, eventualmente, @quantity e @unit.

b) Correccións e adicións existentes no texto orixinal

Para indicar a introdución dun ou varios caracteres que substitúen outros que figuraban no texto empregárase o elemento <subst>.

- Dentro do elemento <subst> o fragmento suprimido marcarase mediante un elemento , seguido dun atributo @type no que se indicará o procedemento empregado para a corrección ("overwrite" para correccións mediante sobreescritura, "overstrike" para correccións mediante riscaduras e borranchos, "cancelado"...).
- O texto engadido indicárase mediante un elemento <add>, seguido dun atributo @place, no que se especificará o lugar en que se insire o texto engadido: "above" (sobre a liña), "below" (baixo a liña), "inline" (na liña), "margin" (na marxe esquerda ou dereita), "top" (na marxe superior da páxina), "bottom" (na marxe inferior da páxina), "overleaf" (no reverso da páxina).

A seguir, ofrécense algúns exemplos de intervención.

b.1. Texto sobreescrito

Trátase de casos en que a corrección se realizou no texto orixinal escribindo un ou varios caracteres directamente enriba dos que se pretendía substituír ou corrixir:

```
é<subst><del  
type="overwrite">b</del><add>v</add></subst>òs para cubizár
```

Se o texto substituído non resultase lexible, empregárase a etiqueta baleira correspondente, pero, sempre que resulte posible, con indicación do número de caracteres eliminados:

é<subst><del type="overwrite" quantity="1" unit="character"/><add>v</add></subst>os.

b.2. Corrección na interliña

Trátase de casos en que a corrección se realizou no documento orixinal riscando un ou varios caracteres e inserindo outros na interliña (superior ou inferior):

Se<subst><del type="overstrike">stanno <add place="above">d'esta</add></subst> non dòu de costas

Tamén neste caso, se o texto riscado non resultase lexible, empregárase a etiqueta baleira con indicación do número de caracteres riscados:

Se<subst><del type="overstrike" quantity="6" unit="character"/><add place="above">d'esta</add></subst> non dòu de costas

b.3. Inserción de texto na liña

Neste caso, a corrección realizouse no orixinal mediante a inserción de texto na propia liña:

non me has <add place="inline">ti</add> de aloumiñar?

Se ademais da inserción no orixinal hai algún elemento riscado, este indicárase mediante un elemento , conforme se indica no apartado a):

<del type="overstrike">ti non me has <add place="inline">ti</add> de aloumiñar?

b.4. Inserción de texto na marxe

Para a inserción de texto na marxe esquerda, dereita, superior ou inferior da páxina, ou no reverso empregárase tamén o elemento `<add>` co atributo `@place`:

```
<add place="margin">palabra</add>
```

```
<add place="top">palabra</add>
```

```
<add place="bottom">palabra</add>
```

```
<add place="overleaf">palabra</add>
```

c) Espazos en branco no texto orixinal

Se o texto presenta un espazo en branco claramente perceptible (como por exemplo o deixado adrede para ser cuberto logo cun dato que se descoñece no momento da redacción), empregárase o elemento `<space/>`, cos atributos `@quantity` e `@unit`:

e viron un `<space quantity="4" unit="character"/>` no ceo

d) Deterioración do soporte

Empregárase o elemento `<damage>` para dar conta daqueles casos en que o soporte material do documento presenta unha área deteriorada, por exemplo, por rotura, queimadura, acción de insectos ou roedores, humidade, etc., de tal xeito que falta unha parte do texto ou ben este resulta total ou parcialmente ilexible. Este elemento poiderá ir acompañado dos seguintes atributos:

- `@agent`, para indicar a causa da deterioración ("fire", "smoke", "mildew" ...)
- `@quantity`, para o número de unidades afectadas ("1", "2", "3" ...)
- `@unit`, para o tipo de unidade ("character", "word", "line" ...)

Así, se, por exemplo, como consecuencia da acción do lume nun documento falta a parte superior dunha páxina, correspondente aproximadamente ás tres

primeiras liñas, indicárase esta circunstancia mediante un elemento baleiro (<damage/>):

```
<damage agent="fire" unit="line" quantity="3"/>
```

Se por esta mesma causa falta a zona esquerda dunha páxina, indicárase ao comezo de cada liña o número aproximado de caracteres que faltan:

```
<damage agent="fire" unit="character" quantity="8"/>
```

Cando dentro dunha área deteriorada hai palabras ou caracteres que se poden ler empregárase un elemento <damage> para delimitar a área. No seu interior, os fragmentos non recuperables indicáranse mediante o elemento <gap> (cos atributos @quantity e @unit). Os fragmentos lexibles indicáranse sen máis, ou, no caso de que a lectura sexa insegura, demarcados por un elemento <unclear>:

```
<damage agent="fire"><gap unit="characters"
quantity="4"/><unclear>eiro</unclear></damage>
```

Cando todo ou parte do texto que falta nunha área deteriorada é reconstruíble de xeito hipotético ou conxectural, este indicárase mediante un elemento <supplied>

```
se          vistas,          meu          <damage
agent="fire"><supplied>amado</supplied></damage>
```

e) Dificultade de lectura e ilexibilidade do texto

Cando o texto presenta nalgún punto dificultades de lectura, empregárase o elemento <unclear> (non demarcado por <damage>) para indicar o fragmento sobre o que editor ten dúbidas:

```
este é o <unclear>verda</unclear>deiro nome
```

f) Notas ou glosas á marxe

Cando o texto presenta glosas ou notas explicativas á marxe, estas indicaranse co elemento <textnote>, que se inserirá no lugar do texto a que corresponda a nota (a continuación da palabra ou palabras glosadas, na fin da liña ou do parágrafo, etc.).

O elemento <textnote> completárase cos atributos @type (para o tipo de información introducida: "gloss"...) e @place (para o lugar en que figura a nota no texto orixinal: "margin", "bottom"...).

```
<l>Embejando o Indo, è o Ganjes</l>  
<l>Os nosos Sar, è Sarela <textnote type="gloss"  
place="margin"><foreign xml:lang="es">Rios de  
Santiago</foreign>.</textnote>.</l>
```

g) Cambios de man

Se nun texto manuscrito aparecen varias mans, os sucesivos cambios indicaranse mediante un elemento <handShift/> acompañado dun atributo @new, no que se indicará a man de que se trate:

```
<handShift new="#h1"/>
```

Cómpre lembrar que as diferentes mans deberán quedar identificadas na cabeceira dentro do elemento <handNote>.

h) Cambios de lingua

Na cabeceira dos textos hai un elemento que permite identificar a lingua principal do texto ("gl") e, ao mesmo tempo, precisar a variedade empregada (galego, galego paródico, galego arcaizante, hibridación galego-castelán, hibridación galego-portugués).

Cando nun texto do corpus aparecen fragmentos escritos noutra lingua, indícase esta circunstancia mediante un elemento `<foreign>` acompañado dun atributo `@xml:lang`, mediante o que se indicará a lingua de que se trate, sempre de acordo co sistema de abreviaturas normalizado (cfr. <http://www.iana.org/assignments/language-subtag-registry/language-subtag-registry>): "es" (español), "la" (latín), "pt" (portugués), etc.

```
<foreign xml:lang="es">SONETO GALLEGO de el Liçenciado  
Pedro Bazquez de Neyra</foreign> <lb/> <foreign  
xml:lang"la">RESPICE FINEM</foreign>
```

i) Imaxes

Cando no texto apareza inserida un debuxo ou imaxe sinalárase a súa presenza mediante o elemento `<figure>`, que poderá conter, á súa vez, outros elementos, coma no exemplo que segue, en que temos un debuxo (indicado mediante un elemento `<graphic>`, que pode ir acompañado dun atributo `@url` en que se indicará o arquivo que o contén) que vai precedido dunha referencia bíblica a xeito de título ou cabeceira (que se marcará mediante un elemento `<head>`) e seguido duns versos (que se indican mediante un elemento `<lg>`):

```
<figure>  
  <head>  
    <p>  
      <foreign xml:lang="la">Frumentum electorum (Zach. 9)</foreign>  
    <p>  
  </head>  
  <graphic url="frumentum.png"/>  
  <lg>  
    <l><foreign xml:lang="es">El pan de los escogidos</foreign></l>  
    <l><foreign xml:lang="es">Corriendo le e de ganar</foreign></l>  
    <l><foreign xml:lang="es">Si acierto bien a parar</foreign></l>  
  </lg>  
</figure>
```