

Modelling the unfinished and fragmentary: the digital archive of the tragedy *The Duke of Parma*

[Um modelo para o inacabado e o fragmentário:
o arquivo digital da tragédia *The Duke of Parma*]

Elena Lombardo*

Keywords

Fernando Pessoa, Scholarly digital edition, *The Duke of Parma*, TEI, Modelling.

Abstract

This article examines the modelling strategies of Fernando Pessoa's *The Duke of Parma* digital archive, a scholarly digital edition of an unfinished English-language tragedy written between c. 1908 and 1935. The play survives in over 180 heterogeneous documents marked by discontinuity, instability, and complex material features. Representing such a corpus digitally poses conceptual and methodological challenges. The TEI-based encoding model prioritises machine readability, interoperability, and minimal interpretation, while addressing material layout, dramatic structure, writing processes, and selected semantic features. Particular attention is given to fragmentariness and Pessoa's compositional habits, both verbal and graphic. The article also considers visualisation choices and technical constraints, advocating an open-ended, scalable approach. The archive is conceived as an "enriched archive" designed to support future editorial work and diverse forms of scholarly research.

Palavras-chave

Fernando Pessoa, Edição filológica digital, *The Duke of Parma*, TEI, Modelação.

Resumo

Este artigo examina as estratégias de modelação do arquivo digital *The Duke of Parma*, uma edição filológica digital de uma tragédia inacabada, em língua inglesa, de Fernando Pessoa. Composta intermitentemente entre 1908 (?) e 1935, a peça sobrevive em mais de 180 documentos heterogêneos, marcados por descontinuidade, instabilidade textual e complexidade material. A sua representação digital coloca desafios conceituais e metodológicos. O modelo de codificação, baseado na TEI, privilegia a legibilidade por máquina, a interoperabilidade e uma intervenção interpretativa mínima, contemplando a disposição material, a estrutura dramática, os processos de escrita e alguns aspetos semânticos. Dá-se especial atenção à fragmentariedade e aos hábitos composicionais do autor, tanto verbais como gráficos. O artigo aborda ainda opções de visualização e constrangimentos técnicos, propondo uma abordagem aberta e escalável. O arquivo é concebido como um "arquivo enriquecido", orientado para futuros desenvolvimentos editoriais e novas formas de investigação.

* University of Lisbon – School of Arts and Humanities, Centre of Linguistics.



The Duke of Parma is an unfinished tragedy in English, written by Fernando Pessoa. Inspired by Shakespearean language, form, and themes,¹ the author worked on this play intermittently for nearly three decades – the longest period he ever dedicated to a single literary project. He began composing it between 1908 and 1909² and continued revising and expanding it until his death in 1935, producing more than 180³ documents now held at the Biblioteca Nacional de Portugal (BNP) in Lisbon.⁴ These features, among others, give the play a unique place in Pessoa's literary production.⁵

As Enrico MARTINES (2024: 3) recalls, although it is only now being published, *The Duke of Parma* has been known to scholars since 1966, when Georg Rudolf Lind mentioned it among Pessoa's English dramas. Subsequent references appeared in the 1970s by Teresa Rita Lopes and Anne Terlinden, with the most recent in Richard Zenith's 2021 biography of Pessoa. An earlier attempt to produce a scholarly edition, led by Carlos Pittella, could not be completed due to lack of funding.

The project *Fernando Pessoa's Duke of Parma: an unpublished Shakespearean tragedy*⁶ represented the first effort to give access to the text of the documents to the public. It aimed at studying and transcribing the documents that make up the dossier of the tragedy and creating a digital archive of the play. The documents were first transcribed by two teams of philologists: Nicolás Barbosa, under the supervision of Jerónimo Pizarro (University of the Andes) and Maria João Sousa, under the supervision of João Dionísio (University of Lisbon). The transcriptions were formalised according to the method developed by the *Equipa Pessoa* for the critical edition of Fernando Pessoa's works in the 1980s and subsequently served as the basis for the text encoding. The resulting archive includes both the facsimiles of the documents

¹ See MARTINES (2024: 2), among others.

² See BARBOSA (2024), who presents a table with the dates of Pessoa's dramatical production, further specifying: "These estimations were first presented in the conference "Entre autos e outros: mapeando o teatro (édito e inédito) de Fernando Pessoa," prepared by Barbosa and Pittella for the *Colóquio Trilogia dos Gigantes* in Lisbon on May 31, 2019, a scholarly event organized by the Centro de Estudos de Teatro da Universidade de Lisboa. These initial estimations are based on the critical editions published as of May 2019 and on archival and inventory work led by Pittella, with the collaboration of Barbosa, at the Portuguese National Library in the summers of 2017 and 2018. Information on some works is still to be determined and, in some cases, the totality of manuscripts, which rest in the Pessoa archive at the Portuguese National Library in Lisbon, remains unclear" (BARBOSA, 2024: 77). Jerónimo PIZARRO (2024), who likewise emphasised that the composition of the play occupied Pessoa for much of his life, has suggested an earlier starting date of 1906-1907.

³ The initially established corpus of known documents (around 170) was publicly revised during the 2023 conference *Fernando Pessoa: The Duke of/in Parma*.

⁴ BNP, Arquivo de Cultura Portuguesa Contemporânea, Espólio 3, envelopes 11⁴D & 11⁵D.

⁵ See, for example, BARBOSA (2024: 69-71), who examines the play's place within the author's dramatic oeuvre, as well as MARTINES (2024), who discusses the motivations behind the tragedy's inception.

⁶ The project's official website can be accessed at: <https://dukeofparma.unipr.it/>. In the following, I will use *DOFP* as the abbreviation for the project.

and their encoded transcriptions, published using Edition Visualization Technology (EVT2 beta2). This paper describes, discusses and seeks to problematise the methodology adopted for the creation of the digital archive by the author of this text.

Defining the object: general modelling principles

When considering *The Duke of Parma* from an editorial perspective, four main features stand out. First, as noted above, it is an unfinished work, composed “in a chaotic and discontinuous manner” (MARTINES, 2024: 6). It survives in a heterogeneous set of documents, ranging from drafts of passages extending over up to nine complete pages (i.e. five leaves with nine inscribed surfaces) to brief fragments consisting of only one or two sentences.⁷ These materials reflect different plot plans and different versions of the text and constitute a dossier that remains highly fragmentary. Second, the documents’ content is “extremely complex [...], written hastily” (MARTINES, 2024: 5), and is often difficult to read – sometimes even indecipherable.⁸ Third, the play is largely unknown and, crucially, had never been edited. Fourth, because the documents are by Fernando Pessoa, they have the potential to attract a broad, international readership. This audience includes, first and foremost, specialists – scholars of Portuguese and English literature, as well as textual critics – but also more general or curious readers. These different readerships are not easily served by a single editorial product: while a critical edition might attempt to address both specialists and curious readers, such interpretative editions ideally presume the prior establishment of a more conservative one.

Against this backdrop, the two-years *DofP* project ultimately positioned itself at the starting point of a possible editorial trajectory – from document to work – and came to define as its main objectives⁹ “to study and transcribe the documents that make up the dossier of the tragedy and [to] create a digital archive of the play”. This entailed deciphering and transcribing the materials and, simultaneously, generating critical knowledge about them. The creation of a digital archive was thus intended to “make all the information on the play’s dossier available for the development of studies of all kinds” (MARTINES, 2024: 6).

⁷ An initial description and a functional classification of the documents was suggested by João Dionísio and Maria João Sousa and at the 2026 *Unveiling the Duke of Parma* conference; the revised version of their lecture can be found in this special issue.

⁸ Jerónimo Pizarro and Nicolás Barbosa addressed the difficulties of transcribing the *Duke* in a communication presented at the aforementioned conference; their lecture can also be found in this special issue.

⁹ Other objectives initially defined were postponed, but the core aim of transcribing the documents and making those transcriptions publicly accessible through a digital platform remained quite constant throughout.

Based on these preliminary considerations, and in order to contextualise the modelling choices involved in the creation of the digital archive, I will begin by examining the concepts of “transcription” and “digital archive” underpinning these objectives.

I will start with a quick commentary about the term “transcription,” recalling an old claim – that every transcription already constitutes an edition: even the act of transferring the text to another medium requires interpretation, since each individual textual feature demands informed decisions. For example, formalising a word written in the interline as an addition already implies a minimal inference about the materiality and temporality of the act. Although editors may explicitly refrain from reconstructing compositional phases, they are nonetheless making assumptions about what they observe on the page.¹⁰

So, “transcribing” is to be intended as “producing a conservative edition” – one involving the minimum necessary inferences. This does not imply avoiding inferences and conjectures altogether, which would be impossible, but rather committing to transparency and limiting interpretations to what can be responsibly pursued within the project’s timeframe. Any subsequent edition, produced with less conservative criteria, should be able to consider this one as its starting point. For example, although the text seems to be primarily in verse,¹¹ the research team is still examining where Pessoa deliberately employed verse, where he turned to prose, and to what extent the perceived irregularities of the verse may stem from editorial interpretation rather than authorial intention.¹² Similarly, indentations of certain words often seem to reflect spatial constraints rather than versification or writing conventions, but this aspect could not be systematically studied. For this reason, one of the criteria was to reproduce the layout of the documents in the edition, transcribing indented words as they appear even when their position was likely dictated by (un)availability of space.

Building on the discussion about the use of the term “transcription”, describing Fernando Pessoa’s *Duke of Parma* Digital Archive as a “digital archive” raises the question of what distinguishes an archive from an edition. This question was widely

¹⁰ These assumptions, as argued by Pizarro and Barbosa in the above-mentioned communication, may even take the form of prior knowledge of specific aspects of the text to be transcribed, thus shaping the transcription itself.

¹¹ Carlos Pittella observed: “However little we know about Pessoa’s three English dramas, they seem to have been written primarily in iambic pentameter—or in Pessoa’s particular version of iambic pentameter, which sometimes allows itself uncanonical stresses (such as on the 3rd syllable, uncanonical in English but perfectly acceptable in Portuguese). Pessoa seems to blend the traditional pentameter with elements of Portuguese decasyllabic verse” (PITTELLA, 2020: 598-599).

¹² This last aspect was addressed and (briefly) discussed among the participants during the 2026 *Unveiling the Duke of Parma* conference.

debated in the early history of digital scholarly editing,¹³ when the appeal of incorporating images encouraged the proliferation of digital editions based on single witnesses, thus reheating the long-standing debate between Lachmannians and Bédierists. In an influential 2019 article (“Edizione documentaria digitale: rinuncia intellettuale o opportunità scientifica?”), Elena Pierazzo argued forcefully that what she termed “a digital documentary edition” – that is, a digital edition based on a single witness or on documents, as opposed to a critical edition, based on many witnesses – is not a passive or intellectually renunciatory enterprise. On the contrary, she defined it as the recording of as many meaningful features of the source document as the editor considers significant, presented in ways judged useful for readers. The emphasis, crucially, is on editorial agency. The selection of features, the modelling of the document, and the modes of visualisation are all products of critical judgement. Today, it is well accepted that an archive would consist merely of images accompanied by metadata. The moment it incorporates encoded transcriptions (encoded conservative editions) it ceases to be a simple repository and actually becomes a documentary digital edition.¹⁴

Should one wish to avoid calling the final product “an edition”, the expression “enriched archive” could provide a suitable alternative, as it preserves the emphasis on aggregation while acknowledging the presence of interpretative and representational layers that exceed mere storage. Such a designation may signal that the resource remains grounded in the logic of the archive – prioritising access, organisation, and documentation – while simultaneously incorporating forms of scholarly intervention that shape how materials are read and understood.

In the context of the present text, however, this terminological distinction is not intended to reopen categorical debates, but rather to clarify usage. The designation “digital archive” was in fact adopted in the project in a broad and pragmatic sense: not as a claim to the absence of editorial agency, but as a description of a resource whose primary orientation is the presentation of documents.

These considerations regarding “transcription” and the notion of “digital archive” informed the modelling of Fernando Pessoa’s *Duke of Parma* Digital Archive. In particular, the emphasis on the documentary character of the materials, the focus on transparency, and the effort of limiting interpretative interventions shaped both the structure of the archive and the encoding strategy. From this perspective, I began by identifying four primary sets of features to be encoded:

¹³ See, among others, PRICE (2009) on the traditional terms “Edition,” “Project,” “Database,” “Archive,” and “Thematic Research Collection” in the context of scholarly digital editions, and ROBINSON (2013) on the relationship between document and work.

¹⁴ Or maybe, adopting terminology more specific to the Portuguese editorial tradition, a “diplomatic-genetic digital edition”.

- Material and layout aspects (blocks of text, headers, lines, etc.).
- Aspects related to the dramatic genre (stage directions and speeches).
- Features of the writing process (additions, deletions, replacements, etc.).
- Selected semantic features: anthroponymy, toponymy and references to the play's structure (acts and scenes).

The encoding model was guided by the following general principles:

- Machine readability was prioritised, favouring highly structured encoding over a more discursive approach, with attributes included wherever possible to enhance clarity and computational accessibility.
- Interoperability was pursued. Substantial shifts from the TEI schema were avoided, with customisation kept within established practices and relying largely on standard elements. Efforts were made to promote openness to future scalability and modification, so that the archive might accommodate later developments without requiring excessive restructuring.
- Emphasis was placed on thorough documentation and transparency of editorial choices.

For the visualisation, the objective was to provide researchers with access to the play for the first time, mirroring the original documents to make them decipherable, without normalisation or regularisation. Regarding the choice of the platform, as will be discussed further below, although all modelling decisions – both general and specific – were independently determined, EVT was adopted as the visualisation platform due to its institutional availability and compatibility with ongoing projects. Nevertheless, it is not the optimal solution for achieving the project's long-term objectives, which is why an open-ended rather than final approach was preferred at all stages.

The *DOFP* encoding manual and the schema, both available in the edition directory on the University of Parma's server,¹⁵ document all decisions and practices adopted. In the next sections, I will comment on some specific encoding aspects, drawing attention to cases that were challenging, need further explanation, or could be subject to debate. The discussion follows the iterative sequence of the work itself: the first application of general criteria, the identification of problems, the definition of new criteria or refinement of previous ones, followed by subsequent applications, and so on.

¹⁵ Accessible via <https://edizionidigitali.unipr.it/evt/data/>

First iteration: challenges of the fragmentary and unfinished

Following from the general principles recalled above, and considering the fragmentary nature of the play, I decided to organise the archive according to material support, with each documentary unit (leaf, fragment, strip, etc.) treated as an autonomous entity. Accordingly, the encoding strategy consists in representing each document individually.

Several considerations informed this decision. First, although many documents can be related to one another (see Dionísio & Sousa's article), the project deliberately refrains from asserting such relationships at the level of representation. Second, methodological coherence required the adoption of a single, consistent unit of work: encoding some XML files at the level of an individual leaf while allowing others to encompass multiple leaves – primarily on the basis of evidence such as authorial pagination – would introduce an uneven degree of interpretation. Third, the choice of the individual document as the basic encoding unit reflects projections regarding future developments towards more interpretative digital editions. Any such future iteration – potentially involving the rearrangement and experimental reading of documents, following models such as the *Arquivo LdoD* (PORTELA & SILVA, 2017), among others – necessarily presumes the individual document as its foundational unit.

To complement this document-level encoding, a “container” XML file (named *dop-main.xml*) was created in order both to provide a higher-level organizational framework for the archive and to accommodate EVT's requirement to process a single XML input.¹⁶ This main document records general information about the project and offers a unified account of material, encoding, and textual aspects of the play, including structural references and onomastic lists. The individual XML files are incorporated into this overarching file through `<xi:include/>` elements.

Given this structure, all occurrences of `@ref` in `<persName>` or other elements intended for retrieval in user searches need to point to the full path of the target element within the container XML. However, such paths might change in the future if the archive is migrated to another visualisation platform. Therefore, to ensure smoother future migrations and maintain consistent cross-referencing, I adopted a system of project-specific prefixes for internal URIs. These prefixes allow elements in individual XML files to be referenced systematically across the archive, without embedding the container file's path directly in each reference. All prefixes are defined in the `<listPrefixDef>` element in the *dop-main.xml* container file. Should the visualisation platform change, it will not be necessary to update every individual `@ref`; adjusting the corresponding XPath expressions in the XSLT will hopefully suffice to maintain correct references.

¹⁶ Encoding all documents in a single file (for example as multiple `<text>` elements) was another option, but it would have undermined the centrality of the document as the basic unit of the archive.

An example of this system in practice can be seen in the encoding of Fernando Pessoa's writing tools. In the `dop-main.xml` container file, each writing tool is defined in `<handDesc>` with a `<handNote>` element, which can then be referenced across the archive using the "hnd" prefix. Below is a `<handNote>` in the `dop-main.xml` file:

```
<handNote medium="blackInk" xml:id="FP-blackInkHandwriting">a black ink pen;</handNote>
```

Example 1. A `<handNote>` element in the `dop-main.xml` file.

In an individual manuscript file, for instance `E3_11-5-D_fl8r-v.xml`, the same tool is referenced using the `@sameAs` attribute and the prefix defined in `<listPrefixDef>`:

```
<handNote medium="blackInk" sameAs="hnd:FP-blackInkHandwriting">Fernando  
Pessoa's handwriting in black ink.</handNote>
```

Example 2. A `<handNote>` element in the `E3_11-5-D_fl8r-v.xml` file.

Similarly, in the `<text>` element, the prefix is used to assign a tool to an insertion or annotation, as in:

```
<add hand="hnd:FP-blackInkHandwriting" place="above">that</add>
```

Example 3. Use of the `@hand` attribute in context.

Following this first strategic decision, the next step was to outline the structure of individual documents. As seen above, a first fundamental modelling criterion was to distinguish between three planes: the material plane of the manuscript or type-written page, the plane of textual content, and the plane of the author's writing process. The material plane concerns the written page as a physical object, including layout, spatial distribution, orientation of writing, and the existence and eventual use of margins; the plane of textual content includes the dramatic text itself; the plane of the writing process includes a full range of authorial actions upon it.

It may be argued that both material and layout features convey information about the writing process and thus belong to the processual dimension. While this observation is, it goes without saying, hardly disputable, the *DOF* model refrained from adopting such a perspective. The encoding records what is visible and/or distinguishable without assigning a chronological status to the phenomena observed.¹⁷ As will be shown in later sections, certain minimal interpretative decisions – such as the reading order of specific additions – have been adopted in the visualisation interface to guide the reader; however, these are encoded separately in dedicated editorial `<note>` elements.

Within this framework, the TEI elements `<facsimile>` and `<text>` were preferred over `<sourceDoc>`. Although `<sourceDoc>` is suited to the representation of

¹⁷ Apart from the minimal inferences I pointed out above when discussing the "transcription" concept.

genetic dossiers, its use would have constrained the explicit encoding of structural and communicative aspects of the dramatic text. In particular, it would not have allowed the representation of theatrical structures alongside the spatial organisation of the page. The combination of <facsimile> and <text> provides instead greater flexibility, making it possible to integrate spatial representation with a controlled encoding of textual structures, while remaining consistent with the semi-diplomatic framework.

Within the <text> element, in the <body>, <div> elements are used to represent the spatial disposition of the text on the page, rather than to express logical or semantic hierarchies within the textual content. In accordance with the modelling principles, in fact, no attempt has been made to impose “structural” hierarchies between the different parts of the text.¹⁸

Distinct <div> elements are employed to mark separate text columns¹⁹ and additions whenever these differ in orientation or in writing tools used. On the contrary, additions sharing the same orientation and written with the same material are treated as belonging to a single column and are therefore encoded within the same <div>, even when they are spatially discontinuous.

Regarding the choice of elements, preference was given, whenever possible, to more general ones. For example, TEI provides two elements for encoding paragraphs: <p>, which marks paragraphs in prose, and <ab>, an anonymous block element. Given that the play is unfinished, the use of <p> would imply the attribution of a conceptual and structural unit — a paragraph — which may not in fact be established in the documents. <ab> was therefore preferred in these cases, as it allows segmentation without presuming a fully defined textual structure. A similar consideration applies to the encoding of lines. As noted above, although in many documents it is possible to identify verse text, the unfinished nature of the play, as well as the presence of certain prose passages, combined with the general criterion of favouring the documents’ topographical representation, led to the decision to use the empty element <lb/> instead of <l>, which is used to mark verses.

These general modelling choices, however, also brought to the fore an issue inherent to the encoding itself. Given the need to privilege the physical appearance

¹⁸ Meaning that, even though an attempt to link scenes not explicitly identified by Pessoa to the plans was carried out by Maria João Sousa, the editorial team didn’t have the time to discuss this subject. The only structures encoded were, then, the ones explicitly signalled by the author.

¹⁹ The term “column” is not used here to denote a fixed or pre-planned layout in the typographic or codicological sense. Rather, it is employed in a looser, descriptive way to refer to coherent blocks of writing that occupy a distinct area of the page and share the same orientation. In this context, “column-type <div>s” reflect the author’s spatial organization of the writing on the page as it appears in the manuscript, without implying intentional layout design, parallel textual organization, or textual hierarchy.

of the page, and the inherently hierarchical nature of XML-TEI, the encoding resulted in several instances of overlapping elements – a well-known problem, particularly acute in the representation of genetic dossiers. These issues were addressed using established TEI-compliant strategies.

```

<pb n="11-4-D-31r" facs="data/images/single/bn-acpe-e-e3-11-4-d-1-100_0065_31_t24-C-R0150.jpg"/>
  <div type="column" facs="#_114D-31r-column">
    <sp who="psn:V">
      <lb break="no"/><speaker><hi rend="underlined"><persName
ref="psn:V">V.</persName></hi></speaker>
      <stage type="delivery">to <persName ref="psn:C3">Cardinal</persName>:</stage>
      <stage type="delivery">(apart)</stage>
      <stage type="setting">-( <persName ref="psn:D1">D</persName> is present)</stage>
    <ab>
      <lb/>How does your <seg type="alt" xml:id="alt_114D-31-1">lordship</seg> <add hand="hnd:FP-
blackInkHandwriting" place="above" xml:id="alt_114D-31-2">eminence</add> find him?
    </ab>
    <sp>
      <alt target="#alt_114D-31-1 #alt_114D-31-2" mode="excl"/>
      <sp who="psn:C3">
        <lb break="no"/><speaker><persName ref="psn:C3"><hi rend="underlined">C.</hi></persName>—
      </speaker>
      <ab>
        Strange, indeed strange. <add hand="hnd:FP-blackInkHandwriting" place="above" xml:id="alt_114D-
31-4">A</add> <seg type="alt" xml:id="alt_114D-31-3">Philosoph<seg><add hand="hnd:FP-blackInkHandwriting"
place="inline" xml:id="alt_114D-31-5"></add>ied<add hand="hnd:FP-blackInkHandwriting" place="inline"
xml:id="alt_114D-31-6"></add></seg></seg> <add hand="hnd:FP-blackInkHandwriting" place="above"
xml:id="alt_114D-31-7">----- er</add>, but <hi rend="circled"><seg type="alt" xml:id="alt_114D-31-9">not
<seg type="alt" xml:id="alt_114D-31-10">in</seg> <add hand="hnd:FP-blackInkHandwriting" place="above"
xml:id="alt_114D-31-11">of</add></seg> <add hand="hnd:FP-blackInkHandwriting" place="below"
xml:id="alt_114D-31-12">outside</add></hi> the schools.
        <lb/><unclear cert="medium">Where<unclear> seemed not life, <hi rend="circled"><seg type="alt"
xml:id="alt_114D-31-13">it seemed to me</seg> <add hand="hnd:FP-blackInkHandwriting" place="above"
xml:id="alt_114D-31-14"><gap reason="illegible"><desc>†</desc></gap>-seemed</add></hi>, <gap
reason="illegible"><desc>†</desc></gap> bitter at men. But I heard
        <lb/>little <subst><del hand="hnd:FP-blackInkHandwriting" rend="strikethrough">of</del> <add
hand="hnd:FP-blackInkHandwriting" place="above">from</add></subst> him. <subst><del hand="hnd:FP-
blackInkHandwriting" rend="strikethrough">But</del> <add hand="hnd:FP-blackInkHandwriting"
place="above">Nevertheless</add></subst>, I have not <seg xml:id="seg_114D-31-1" rend="hesitationMark"><gap
reason="illegible"><desc>†</desc></gap> with <gap reason="illegible"><desc>†</desc></gap></seg>.
        <metamark function="dissatisfaction" target="#seg_114D-31-1"/>
      </ab>
    </sp>
    <join type="alt" target="#alt_114D-31-4 #alt_114D-31-5 #alt_114D-31-6 #alt_114D-31-7"
result="alt_114D-31-8" xml:id="alt_114D-31-8"/>
    <alt target="#alt_114D-31-3 #alt_114D-31-8" mode="excl"/>
    <alt target="#alt_114D-31-10 #alt_114D-31-11" mode="excl"/>
    <alt target="#alt_114D-31-9 #alt_114D-31-12" mode="excl"/>
    <alt target="#alt_114D-31-13 #alt_114D-31-14" mode="excl"/>
    <lb/><metamark function="separation" rend="halfRule centred">__</metamark>
  </div>

```

Example 4. Use of the <join/> element in context (11⁴ D-31^r).

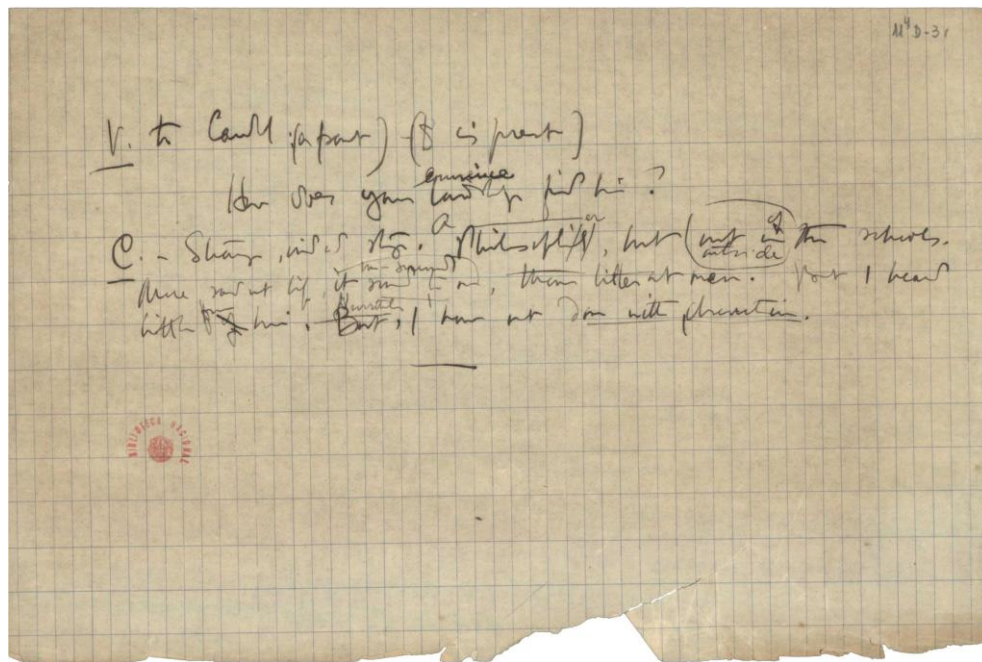


Fig. 1. (BNP/E3, 114D-31).

In particular, in cases where a single speech is materially fragmented across different structural divisions – such as recto and verso surfaces or differently oriented marginal additions – its @xml:id sequential number²⁰ is extended with a letter suffix to distinguish the individual segments. These segments are not intended to represent distinct speeches, but rather discontinuous material manifestations of a single utterance. Their conceptual unity is therefore restored with the <join/> empty element, which explicitly links the relevant <sp> elements and allows them to be treated as a single speech while preserving their distinct material and spatial contexts.

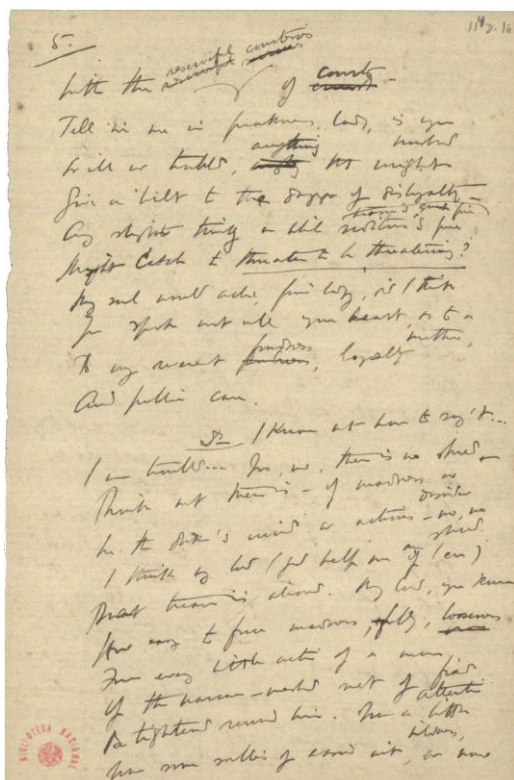
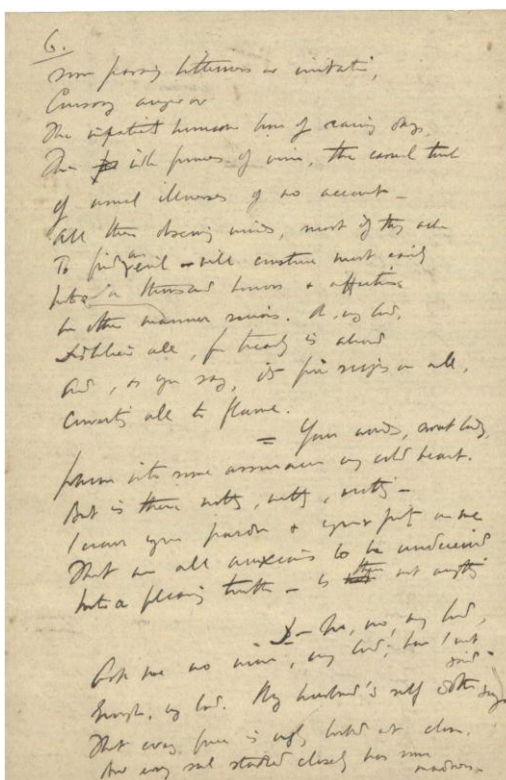
```
<sp who="psn:D2" xml:id="sp_114D-16-2a">
  <lb break="no"/><speaker><hi rend="indented underlined"><persName ref="psn:D2">D<hi
rend="superscript">s</hi></persName> </hi></speaker>
  <ab>
    <lb>I know not how to say it...
    <lb>I am troubled... No, no, there is no shred —
    <lb>Think not there is — of madness or
    <lb><hi rend="indented">disorder</hi>
    <lb>In the duke's mind or actions — no, no
    <lb><hi rend="indented">shred.</hi>
    <lb>I think my lord (God help me <seg type="alt" xml:id="alt_114D-16-3">if</seg> <add
hand="hnd:FP-blackInkHandwriting" place="above" xml:id="alt_114D-16-4"><unclear
cert="medium">once</unclear></add> I err)
    <lb>That treason is <unclear cert="medium">abroad</unclear>. My lord, you know
    <lb>How easy to force madness, folly, <subst><del hand="hnd:FP-blackInkHandwriting"
rend="strikethrough"><gap reason="illegible"><desc>†</desc></gap></del> <add hand="hnd:FP-
blackInkHandwriting" place="above">looseness</add></subst>
    <lb>From every little action of a man
    <lb>Of the <unclear cert="medium">narrow-meshed</unclear> not of fixed attention
    <lb>Be tightened round him. Nor a little
```

²⁰ Each <sp> element, encoding a speech, carries, along with other attributes, an @xml:id with the form “sp_shelfmark-ASequentialNumber”.

```

</lb><hi rend="indented">wildness,</hi>
</lb>Nor some sullies of easier wit, or none
</ab>
</sp>
<alt target="#alt_114D-16-3 #alt_114D-16-4" mode="excl"/>
</div>
<pb n="11-4-D-16v" facs="data/images/single/bn-acpc-e-e3-11-4-d-1-100_0036_16v_t24-C-R0150.jpg"/>
<div type="number" facs="#_114D-16v-number">
<ab><num><hi rend="underlined">6.</hi></num></ab>
</div>
<div type="textColumn" facs="#_114D-16v-column">
<sp who="psn:D2" xml:id="sp_114D-16-2b">
<ab>
<lb>Some passing bitterness or imitation,
<lb>Cursory anger or <space quantity="20" unit="char"><desc>□</desc></space>
<lb>The impatient humour born of raring days,
<lb>This <del hand="hnd:FP-blackInkHandwriting" rend="strikethrough"><gap
reason="illegible"><desc>†</desc></gap></del> idle pureness of mine, the casual tomb
<lb>Of usual illness of no account —
<lb>All these obsessing minds, must if they ache
<lb>To find <add hand="hnd:FP-blackInkHandwriting" place="above">an</add> evil — will <unclear
cert="medium">costume</unclear> most easily
<lb>Into a thousand horrors & affections
<lb>In other manner <unclear cert="medium">serious</unclear>. Oh, my lord,
<lb>Disbelieve all, for treachery is <unclear cert="medium">abroad</unclear>
<lb>And, as you say, its fire <unclear cert="medium">swings</unclear> on all,
<lb>Converting all to flame.
</ab>
...
<join type="sp" target="#sp_114D-16-2a #sp_114D-16-2b" result="sp_114D-16-2" xml:id="sp_114D-16-2"/>

```

Example 5. Use of the <join/> element in context (11⁴D-16^v).Fig. 2. (BNP/E3, 11⁴D-16^r).Fig. 3. (BNP/E3, 11⁴D-16^v).

When structural divisions overlap additions, the empty `<addSpan/>` element is used. It is placed at the beginning of the addition and carries a `@spanTo` attribute whose value points to the `@xml:id` of an empty `<anchor/>` element positioned at the point where the addition ends. An example is provided below.

```
<sp>
  <ab>
    <lb/><addSpan hand="hnd:FP-graphitePencil" place="inline" spanTo="#pencilEnd"/><hi
rend="indented">— What's to be done</hi>
  </ab>
</sp>
<sp>
  <ab>
    <lb/>— This: he must be deposed.
  </ab>
</sp>
<sp>
  <ab>
    <lb/>— <hi rend="halfCircledAbove"><seg type="alt" xml:id="alt_115D-29-7">Why, sir, he
strives</seg> <add hand="hnd:FP-graphitePencil" place="above" xml:id="alt_115D-29-8">What <unclear
cert="medium">shows</unclear> he</add></hi>
    <lb/>To be <subst><del hand="hnd:FP-blackInkHandwriting" rend="overwritten">des<anchor
xml:id="pencilEnd"/></del><add hand="hnd:FP-blackInkHandwriting" place="inline">dep</add></subst>ised?
Careless of state he is not,
    <lb/>Not easily <add hand="hnd:FP-blackInkHandwriting" place="above">to be</add> flattered into
countenance
    <lb/>Of enemies or <space quantity="15" unit="char"><desc>□</desc></space> to the dukedom,
    <lb/>No other <subst><del hand="hnd:FP-blackInkHandwriting" rend="overwritten">s</del><add
hand="hnd:FP-blackInkHandwriting" place="inline">S</add></subst>upervisor nor approver
  </ab>
```

Example 6. Use of `<addSpan/>` in context.

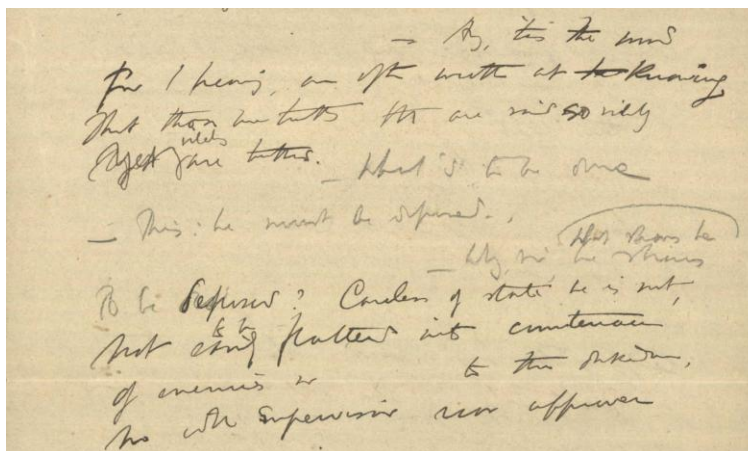


Fig. 4. (BNP/E3, 11⁵ D-29^v; detail).

Finally, when substitutions span multiple elements, the empty `<substJoin/>` element is employed. In such cases, an `@xml:id` attribute is assigned to both the `<del>` and the corresponding `<add>` elements, and a `<substJoin/>` element is then used with a `@target` attribute pointing to their respective `@xml:id` values.

```

<div type="addition" facs="#_115D-5v-additionB">
  <sp xml:id="sp_115D-5-1d" who="psn:D1">
    <ab>
      <lb/>comes. Take me in to the stars.<anchor xml:id="verticalAddEnd"/><add hand="hnd:FP-
blackInkHandwriting" place="inline">We need a
      <lb/><unclear cert="low">great</unclear> purpose in the universe, but <del hand="hnd:FP-
blackInkHandwriting" rend="strikethrough" xml:id="substCircAdd">we tire of talking</del></add>
    </ab>
  </sp>
  <note type="editorial" resp="prj:EL">These lines were added above the initial text of the addition as a result of
limited available space.</note>
</div>
<div type="addition" facs="#_115D-5v-additionC" rend="right-aligned">
  <sp xml:id="sp_115D-5-1e" who="psn:D1">
    <ab>
      <lb/><add hand="hnd:FP-blackInkHandwriting" place="top" rend="circled" xml:id="circAdd"><metamark
function="linking"/>there is too
      <lb/>much speak-<lb break="no"/>ing already.</add></ab>
    </sp>
  </div>
  <substJoin target="#substCircAdd #circAdd"/>

```

Example 7. The use of the <substJoin/> empty element in context (BNP/E3, 11⁵ D-5^v).

So far, the discussion has focused on two of the four sets of features identified in the initial modelling phase. A further set concerns semantic features, including characters, places, and the structure of the play. In this area, too, the unfinished nature of the work gave rise to a number of specific challenges.

Pessoa identified his characters using both abbreviations and full names, both specific forenames or generic role, or noble titles. Moreover, he changed the forename of a character during the writing process – a figure referred to by the initial “V.” appears both as “Vincenzo” (for example, BNP/E3, 11⁵ D, fol. 43)²¹ and as “Valerio” (for example, fol. 11⁵ D-60). In addition, the author planned characters in his plans and lists that were never fully realised in the dramatic text itself, as some figures mentioned there do not appear in the documents encoded within the project. Further complexity arises from the fact that Pessoa occasionally wrote dialogue without explicitly indicating the speaker. While in many cases a simple dash marks the beginning of a speech, for at least one character he uses the initial “X”, which may function as a kind of placeholder. As a result, it was necessary to allow for multiple references to the same character, as well as to account for varying degrees of certainty regarding the existence and identification of individual *dramatis personae*.

For this reason, each potentially identifiable character was encoded as a <person> element and assigned a unique @xml:id in the <listPerson> under the <sourceDesc> element in the dop-main.xml file. The identifier consists of the initial letter of the most frequently used name, followed by a number where more than one character shares the same initial. The numbering has no interpretative value and reflects solely the order in which the characters were retrieved from the sources.

²¹ From this point onwards, for brevity, the first part of the shelfmark will be omitted in references to the documents (e.g.: 11⁴ D-1).

Each `<person>` element contains a `<persName>`, which may include a `<name>`, `<forename>`, `<roleName>`, or other appropriate name components, depending on the specific character. As for `<roleName>` elements, a further `@type` attribute is specified. Possible values include:

- office (e.g.: minister),
- profession (e.g., nurse),
- nobility (e.g.: Duchess),
- religious (e.g.: bishop),
- honorific (e.g.: `<addName type="honorific"><roleName type="nobility">Count</roleName> Alberto</addName>`),
- relational (e.g.: A's son).

When Pessoa indicated both the full name and the initial, the latter was encoded using the `@full` attribute, with the value "init" (e.g.: `<addName type="forename" full="init">A</addName>`).

Additional information – such as alternative names, abbreviations, titles, modes of reference and discussion about potential interpretative problems – is encoded discursively within a `<note>`.²²

The section is completed by a list of relationships between characters, encoded by means of `<relation>` elements grouped within a `<listRelation>` element. Each relation is qualified by a set of attributes: `@type`, whose values distinguish between personal and professional relationships (e.g. "personal", "social"); `@name`, which specifies the nature of the relationship (such as *spouse*, *brother*, *parent*, or *maid*); and attributes indicating mutuality or directionality. Symmetrical relationships are encoded using `@mutual`, while asymmetrical ones employ `@active` and `@passive`, each pointing to the relevant character identifiers.

```
<person xml:id="A1">
  <persName>
    <forename>Alberto</forename>
  </persName>
  <note>May also appear in abbreviated form as "<addName type="forename"
full="init">A</addName>",
    or be referred to by his title ("<addName type="honorific"><roleName
type="nobility">Count</roleName> Alberto</addName>"),
    or by his office ("<roleName type="office">Chief Chancellor</roleName> – or "<roleName
type="office">first Chancellor</roleName>" – both abbreviated as
    "<addName type="office" full="init">C</addName>"; <hi rend="italics">cf</hi>. fols. <locus
n="114D-21v">11<hi rend="superscript">4</hi>D-21v</locus> and <locus n="115D-60r">11<hi
rend="superscript">5</hi>D-60r</locus>). A comparison between plot plans (fols. <locus n="114D-2">11<hi
rend="superscript">4</hi>D-2</locus>, <locus n="115D-43">11<hi rend="superscript">5</hi>D-43</locus>, <locus
n="115D-44">11<hi rend="superscript">5</hi>D-44</locus> and <locus n="115D-80">11<hi
rend="superscript">5</hi>D-80</locus>) and the content of the documents shows that the designation "<addName
type="office">Minister</addName>", often abbreviated as
```

²² The mapping of each initial to the corresponding full name was suggested by Maria João Sousa and revised by the other researchers of the *DOFP* project.

"<addName type="office" full="init" cert="medium">M</addName>", can be also associated with Alberto. This identification appears to be plausible in several instances (e.g.: fols. <locus n="114D-6r">11<hi rend="superscript">4</hi>D-6r</locus>, <locus n="11-4-D-24r">11<hi rend="superscript">4</hi>D-24r</locus>, etc.); however, in fol. <locus n="114D-3av">11<hi rend="superscript">4</hi>D-3av</locus>, a character identified by the initial "M" is shown speaking to a character identified by the initial "A". This suggests that, at least in this case, the initial "M" may not refer to Alberto. For this reason, the present list also includes a separate character identified by the initial "M", in the expectation that further research may clarify the relationship between these references.

<lb/></note>
</person>

Example 8. A <person> element in the <sourceDesc>.

Just as the *dramatis personae*, the structure of the play is likewise unstable. The existence of at least four distinct structure plans (on fols. 11⁴D-2, 11⁵D-43, 11⁵D-44 and 11⁵D-80) suggests the play's structure evolved over time.²³ For this reason, the encoding followed the plan recorded on fol. 11⁴D-2, which seems to be the more complete. In doing so, the purpose is not to establish a definitive ordering, but rather to create stable identifiers that can be linked to specific occurrences of structural labels and references across the documents. This approach allows such information to be retrieved more easily and analysed if needed in future research on the play.

Each act is represented by an <item> element containing a nested list of scene items, each identified by a <label> indicating the scene number. In the version encoded, the outline comprises five acts: Act I includes seven scenes, Act II five scenes, Act III seven scenes, Act IV six scenes, and Act V eight scenes.

Second iteration: a model for the author's compositional self-dialogue?²⁴

However, as the encoding advanced within the framework outlined above, it became evident that the four initial dimensions were, at times, more complex than I had originally anticipated. In particular, the *Process of writing* category appeared to be broader and richer than expected.

Consider structure labels as: "Act I" or "Act I Sc I". Initially, I encoded them with <head>, which "contains any type of heading" (TEI, 4.2.1). Yet, as I continued working through the transcriptions, I encountered other expressions – like "Scene (b)", "Penultimate scene", "Speech of religion", "Other Scene" – that called this choice into question. These instances suggested that I might be attempting to stabilise unstable or provisional references by assigning them to a category that was not entirely appropriate. In fact, if I had opted for <ab> rather than <p> on the grounds that the use of the latter would imply a conceptual and structural unit not necessarily

²³ A comparison of the plot plans was carried out by Maria João Sousa within the *DOFP* project.

²⁴ This section was thoroughly discussed during the *Unveiling the Duke of Parma* conference. In relation to the work presented there, I have made further adjustments to the model, following observations by Elli Bleeker, to whom I am very grateful for the engaging and constructive dialogue, not only during the presentation but throughout our whole stay in Parma.

established in the play, then the use of <head>, a structurally specific element, could be equally problematic.

Given that one of the aims of the encoding was to enable the retrieval of all occurrences indicating the presence of a developing structure within the play, and considering that anthroponyms and toponyms were being encoded by means of <persName> and <placeName>, I explored the possibility of adopting <rs> for these instances. This element, defined as a “referencing string” that “contains a general purpose name or referring string” (TEI, 14.2.1 and 3.6.1), could be combined with an attribute @type=“structural”, as shown in the following table.

Attribute	Possible values	Description	Additional attribute of the first	Possible values	Description	Additional attribute of the first	Possible values
@type (mandatory)	structural	Marks authorial indications related to the structure of the work (acts, scenes, sections)	@subtype (mandatory)	defined	Refers to an explicitly identified structural unit	@ref (mandatory)	pointing to the dop-main.xml file by means of the prefix “str”
				generic	Indicates a vague or provisional structural placeholder		

Table 1. Attributes for structure labels.

Once the use of <head> for such instances was excluded, I began to question its application in other cases, such as for expressions like “Duke of Parma”, “D. of P.”, occurring at the top of certain documents. Although these expressions precede passages belonging to the fictional text, they do not consistently function as part of that text. Rather, they appear to act as orienting markers, signalling to Pessoa that the text which follows pertains to the *Duke of Parma* and not to another work on which he could be working concurrently.²⁵

Since these occurrences had initially been encoded with <head>, I first experimented with reclassifying them as <rs> elements, introducing a new attribute (@inclusive) to distinguish them from the “structural” markers above.

However, this solution proved only partially satisfactory. The issue was further discussed at the *Unveiling the Duke of Parma* conference, where Elli Bleeker suggested the use of <label>, which “contains any label or heading used to identify part of a text, typically but not exclusively in a list or glossary” (TEI, 3.8). Following this suggestion, both act titles and instances of the play’s title functioning as inclusive or orienting references were encoded using the <label> element, retaining the set of attributes previously defined. Illustrative examples in context are provided below.

²⁵ The same hypothesis was independently developed by Dionísio & Sousa and presented at the *Unveiling the Duke of Parma* conference.

Note that this applies to cases where the author marks the association of a passage with the work as a whole, such as when writing the title of the play. In some instances, however, ambiguity arises between a reference to the Duke's name as a character and as part of the title of the work. In such cases, encoding decisions were guided by contextual evidence and by the apparent function of the label within the writing process.

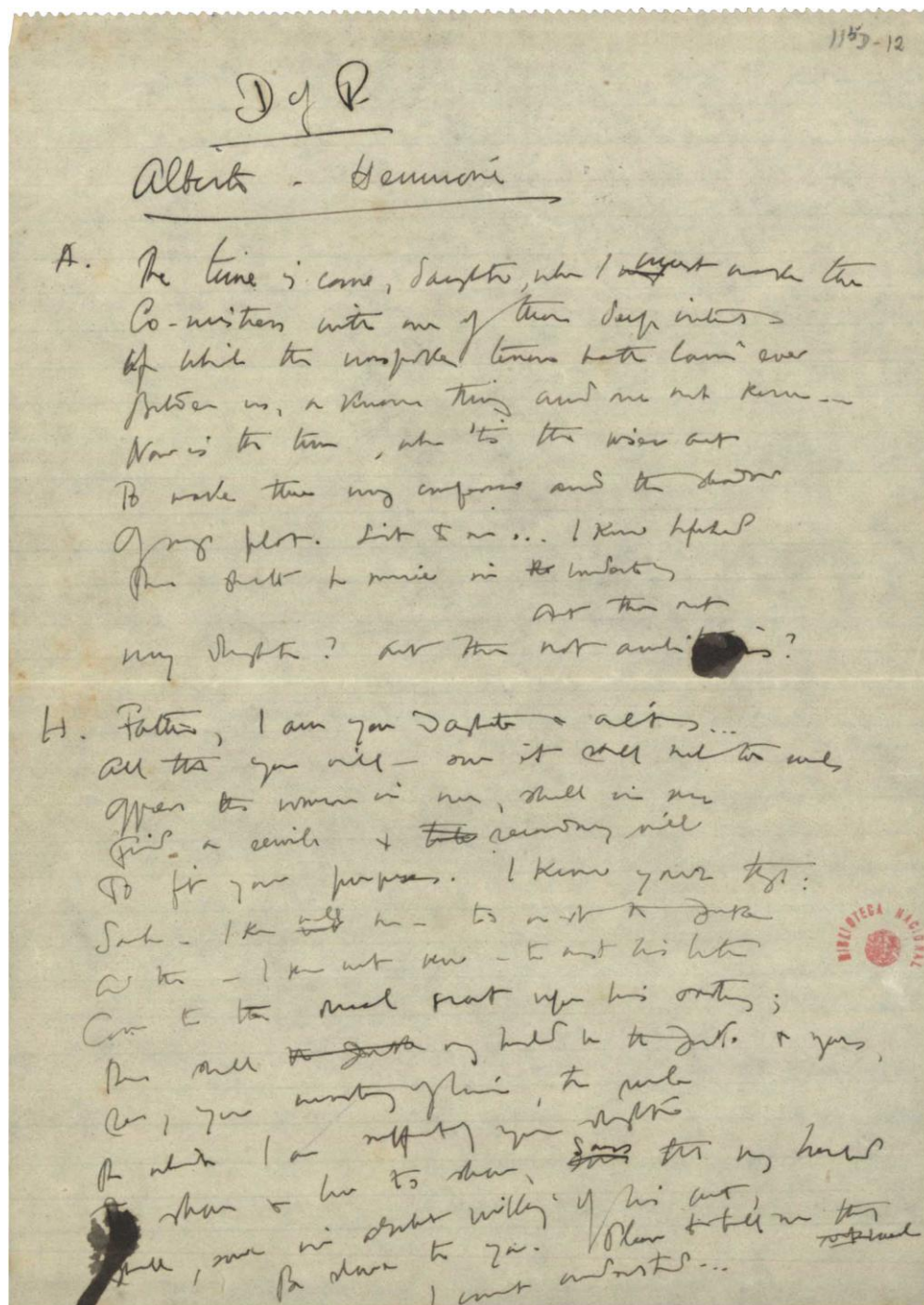


Fig. 5. Example of a "D. of P." header (BNP/E3, 11⁵ D-12⁹).

```
<ab rend="centred">
  <hi rend="underlined"><label type="structural" subtype="defined" ref="str:ActIV">Act
IV</label></hi>
  <lb/><hi rend="underlined"><stage type="setting">Council-Hall</stage></hi>
</ab>
```

Example 9. Use of <label type="structural" subtype="defined"> in context.

```
<ab rend="centred"><hi rend="underlined"><label type="structural" subtype="generic">Other
Scene.</label></hi></ab>
  <sp>
    <p>
      <lb/><del hand="hnd:FP-blackInkHandwriting" rend="strikethrough">This</del> What is this man? Do I
know him
      <lb/><hi rend="indented">aught?</hi>
    ...
```

Example 10. Use of <label type="structural" subtype="generic"> in context.

```
<ab rend="centred"><hi rend="underlined"><label type="structural" subtype="generic">Monologue just before
murder</label></hi></ab>
  <sp who="psn:A1">
```

Example 11. Another example of <label type="structural" subtype="generic"> in context.

```
<ab rend="centred"><hi rend="underlined circled"><label type="structural"
subtype="generic">Moon</label></hi></ab>
  <sp who="psn:D2">
    <lb break="no"/><speaker><hi rend="indented"><hi rend="underlined"><persName
ref="psn:D2">Duchess</persName></hi></hi></speaker>
    <ab>
      it is, it is, <subst><del hand="hnd:FP-blackInkHandwriting" rend="strikethrough">my lord</del>
<add hand="hnd:FP-blueInkHandwriting" place="above">my love</add></subst>
    </ab>
  </sp>
```

Example 12. A third example of <label type="structural" subtype="generic"> in context.

```
<ab><hi rend="underlined"><label type="inclusive">Duke of Parma</label></hi></ab>
<sp>
  <ab>
    <lb/>Thou art a fool, who borrowest
    <lb/>sincerity of thine imagination, which
    <lb/>in turn thou borrowest <add hand="hnd:FP-graphitePencilHandwriting" place="above">is a
name</add> of thy folly,
    <lb/>making a shadow of thyself <unclear cert="medium">lie</unclear> to <seg type="alt"
xml:id="alt_114D-4-1">an</seg> <add hand="hnd:FP-graphitePencilHandwriting" place="above" xml:id="alt_114D-
4-2">the</add>
    <lb/>other <seg type="alt" xml:id="alt_114D-4-3">shadow</seg> <add hand="hnd:FP-
graphitePencilHandwriting" place="above" xml:id="alt_114D-4-4">substance</add> of <seg type="alt"
xml:id="alt_114D-4-5">thyself</seg> <add hand="hnd:FP-blueInkHandwriting" place="above" xml:id="alt_114D-4-
6">thy folly</add>,
  </ab>
</sp>
```

Example 13. Use of <label type="inclusive"> in context.

As the encoding further progressed, I began to encounter clearer instances of “meta-text” – segments not belonging to the narrative layer, instead functioning as forms of authorial self-communication. Examples include formulations such as: “(these phrases begin the reconciliation)” (11⁴D-8^r), “good” (11⁵D-36^r), or “Explanatory: The Duke’s constant sexual allusions and their purpose. They are not the author’s; they belong to the author’s conception of the character of the Duke” (11⁵D-83^r).

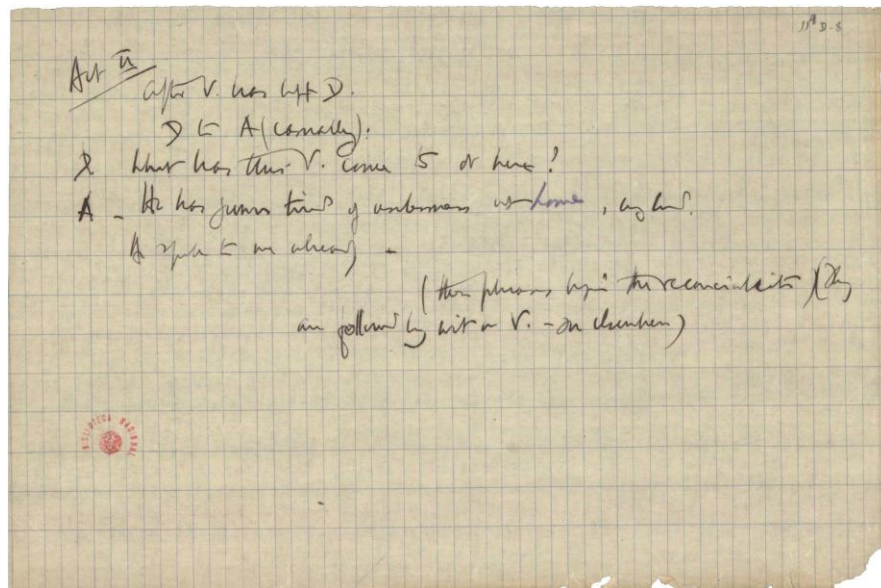


Fig. 6. A meta-textual instance (BNP/E3, 11⁴D-8^r).

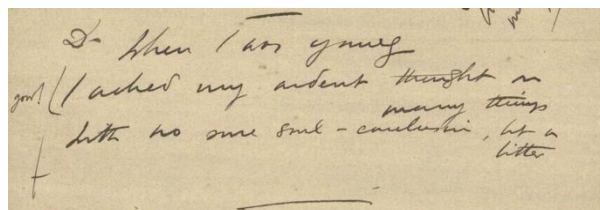


Fig. 7. A meta-textual instance, "good!" (BNP/E3, 11⁵D-36^r; detail).

Such passages do not contribute to the fictional world of the play. Rather, they mark moments in which the author appears to step outside the dramatic discourse and reflect upon the text in progress. In doing so, they articulate a provisional organisation of the material, indicate how the work is being conceptualised, and may signal distinct phases of composition.

Within this model, I refer to these phenomena as meta-text. This category encompasses those segments in which the author reflects upon, evaluates, or directs the ongoing process of writing. It thus captures a range of functions, including internal structuring cues, reminders, evaluative remarks, and more extended explanatory comments. Taken together, these elements document the author's evolving negotiation with structure, meaning, and trajectory, and provide insight into compositional thinking as it unfolds.

In an initial phase, these segments were likewise encoded using <rs>, progressively extending its use to cover all instances in which the author stepped outside the narrative frame. Although the notion of "reference" seemed at first ill-suited to this context, its breadth could be invoked to justify such usage: what mattered was not whether a string pointed to an external entity, but that it functioned as a semantically distinct marker whose meaning could not be reduced to its literal wording alone. From this perspective, meta-text was understood as consisting precisely of

such strings – elements that registered operations of structuring, evaluation, or planning, rather than contributing directly to the fictional or dramatic content.

From an analytical and querying perspective, the decision to encode all these instances under a single element offered certain advantages, as it enabled the identification of such moments through a single access point. At the same time, however, this approach risked introducing a degree of semantic flattening, as different functional types were subsumed under the same element. A controlled set of attributes was therefore developed to differentiate between them and to retain a minimal level of interpretive nuance, as described in **Error! Reference source not found..**

Even so, this solution remained conceptually unsatisfactory. The justification underpinning the use of <rs> felt strained and ultimately insufficiently precise. Following further reflection (and discussion at the *Unveiling the Duke of Parma* conference), I eventually adopted the <note> element for the encoding of these segments, as suggested by Elli Bleeker. Defined as an element that “contains a note or annotation” (TEI, 3.9.1), <note> offers a more appropriate framework for capturing material that is, in some sense, set apart from the main textual flow, even if it does not fully resolve the underlying problem.

The TEI Guidelines describe <note> primarily in terms of its material and positional features:

The <note> element is provided for the encoding of discursive notes, whether already present in the copy text or supplied by the encoder [...]. A note is any additional comment found in a text, marked in some way as being out of the main textual stream. All notes should be marked using the same tag, note, whether they appear as block notes in the main text area, at the foot of the page, at the end of the chapter or volume, in the margin, or in some other place.

Notes may be in a different hand or typeface, may be authorial or editorial, and may have been added later.

(<https://www.tei-c.org/release/doc/tei-p5-doc/en/html/CO.html#CONONO>)

While this definition foregrounds visual and documentary aspects rather than semantic function, it nevertheless provides a useful point of departure. In particular, the notion of textual material that stands apart from the main discourse resonates with the status of compositional self-dialogue, which does not belong to the fictional layer of the work. However, the correspondence is not exact. These segments are not merely “additional comments”, nor are they always clearly separable from the unfolding of the text in spatial or typographic terms. Rather, they are integral to the writing process itself. The use of <note> should therefore be understood as a pragmatic modelling decision: it provides a flexible and TEI-compliant means of representing material that operates alongside, but not within, the fictional discourse, even if it does not fully capture its functional specificity. Some examples follow.

Attribute	Possible values	Description	Additional attribute of the first	Possible values	Description	Additional attribute of the first	Possible values
@type (mandatory)	synoptic	Provides a descriptive summary of the content of an adjacent or preceding structural unit	N/A	N/A	N/A	N/A	N/A
	plot-oriented	Encodes notes related to plot planning, character behaviour, or narrative development					
	choice	Encodes words used to indicate alternative readings					
@type (mandatory)	assessing	Encodes authorial evaluation, judgement, or problem-flagging	@subtype (mandatory)	approval	Signals satisfaction or endorsement	@target (optional) @rend (optional)	Pointing to the @xml:id of the correspondent <seg>
				dissatisfaction	Signals perceived inadequacy		
				disclaimer	Clarifies authorial intent or distances the author from a textual effect		
				rejection	Marks complete rejection or negation of a passage or idea		
@type (mandatory)	procedural	Encodes prompts that accompany and externalise the act of writing itself, including cues for continuation, decision-making, etc.	@subtype (mandatory)	completion	Signals deferred textual production or improvement	N/A	N/A
				navigating	Encodes instructions for reading (or “navigating”) the text expressed in words		
@hand (optional)			N/A	N/A	N/A	N/A	N/A

Table 2. Attributes for meta-text, ultimately encoded with the <note> element.

```

<ab rend="centred"><hi rend="underlined"><label type="structural" subtype="defined"
ref="str:ActV">V</label></hi>
<lb/><note type="synoptic"><hi rend="underlined"><persName ref="psn:B2">Bishop</persName>'s
conflict</hi>
<lb/>to <persName ref="psn:V">Valerio</persName> (& discussion)</note></ab>

```

Example 14. Use of <note type="synoptic"> in context.

```

<ab><lb/>
<note type="plot-oriented">When <persName ref="psn:D1">Duke</persName> appears, he speaks very
<lb/>naturally & normally (calculatingly,
<lb/>of course), even accentuating, with
<lb/>a subliminal irony his love of
<lb/>custom & appearance.
</note>
</ab>

```

Example 15. Use of <note type="plot-oriented"> in context.

```
<ab>
    <note type="plot-oriented"><hi rend="indented">(these phrases begin the reconciliation)(They</hi>
    <lb/>are followed by wit <unclear cert="medium">on</unclear> <persName ref="psn:V">V.</persName>
    — <unclear cert="medium">Or</unclear> elsewhere)</note>
</ab>
```

Example 16. Another example of <note type="plot-oriented"> in context.

```
<add hand="hnd:FP-blackInkHandwriting" place="rightMargin" rend="writingMode(verticalLR)
hesitationMark"><lb/>Perchance he of himself will go to madness
    <lb/>And thus my curse is traced by him. <note type="assessing" subtype="dissatisfaction">(<unclear
cert="medium">garble</unclear>)</note></add>
```

Example 17. Use of <note type="assessing" subtype="dissatisfaction"> in context.

```
<add hand="hnd:FP-blackInkHandwriting" place="margin" rend="marginMetamark"><note type="assessing"
subtype="approval" target="#seg_115D-36-1">good!</note></add>
```

Example 18. Use of <note type="assessing" subtype="approval"> in context.

```
<ab>
    <lb/><note type="assessing" subtype="disclaimer">(Explanatory: <persName ref="psn:D1">The
Duke</persName>'s
    <lb/>constant sexual allusions
    <lb/>and their <hi rend="underlined">purpose</hi>. They
    <lb/>are not the author's; they
    <lb/>belong to the author's
    <lb/>conception of the
    <lb/>character of <persName ref="psn:D1">the Duke</persName>.</note>
</ab>
```

Example 19. Use of <note type="assessing" subtype="disclaimer"> in context.

```
<lb/>Lot spread the report of <persName ref="psn:D1">the Duke</persName>'s
    <lb/>madness; at the same time <persName ref="psn:A1"><add hand="hnd:FP-blackInkHandwriting"
place="above">lord</add> Alberto</persName>, the
    <lb/><seg type="alt" xml:id="alt_114D-21-3"><persName ref="psn:A1">Chief
Chancellor</persName></seg> <add hand="hnd:FP-blackInkHandwriting" place="above" xml:id="alt_114D-21-
4"><persName ref="psn:A1">first counsellor</persName></add> attempts to <del hand="hnd:FP-
blackInkHandwriting" rend="strikethrough"><gap reason="illegible"><desc>†</desc></gap></del> do
    <lb/>the same thing in another way — <hi rend="bracketed"><seg type="alt" xml:id="alt_114D-21-5">to
present a <del hand="hnd:FP-blackInkHandwriting" rend="strikethrough">some</del> child of the <persName
ref="psn:D1">Duke</persName> <add hand="hnd:FP-blackInkHandwriting" place="below">so he might dominate
completely</add></seg></hi> <hi rend="circled"><note type="assessing" subtype="choice">or</note></hi> <hi
rend="bracketed"><seg type="alt" xml:id="alt_114D-21-6">to let the madness be confined, then to <seg type="alt"
xml:id="alt_114D-21-7b">overwhelm <persName ref="psn:G">Girolamo</persName> & his lot with the proof
of <persName ref="psn:G">Girolamo</persName>'s <hi rend="underlined">illegitimacy</hi></seg></seg></hi> <hi
rend="circled"><note type="assessing" subtype="choice">or</note></hi> <hi rend="openingBracket"><seg
type="alt" xml:id="alt_114D-21-7a">to</metamark function="linking" target="#alt_114D-21-7b"/></seg></hi>
    <lb/><metamark function="separation" rend="halfRule centred">__</metamark>
</ab>
<join type="alt" target="#alt_114D-21-7a #alt_114D-21-7b" result="alt_114D-21-7" xml:id="alt_114D-21-
7"/>
<alt target="#alt_114D-21-3 #alt_114D-21-4" mode="excl"/>
<alt target="#alt_114D-21-5 #alt_114D-21-6 #alt_114D-21-7" mode="excl"/>
</div>
```

Example 20. Use of <note type="assessing" subtype="choice"> in context.

```
<ab> <add hand="hnd:FP-blackInkHandwriting" place="bottom" rend="complexUpsideDown"><hi rend="underlined
overlined"><note type="procedural"
subtype="choice"><choice><abbr>Pa</abbr><expan>P<ex>ar</ex>a</expan></choice> escolher</note></hi></add>
</ab>
```

Example 21. Use of <note type="procedural" subtype="choice"> in context.

```
<note type="navigating"><hi rend="underlined">Backward</hi></note>
```

Example 22. A <note function="navigating"> element.

While <note> elements provide a means to encode authorial self-communication expressed through language, the documents also contain another layer of meta-textual activity: marks, symbols, and visual cues that the author uses to manage, manipulate, or emphasise the text. These signs – numbers, arrows, crosses, underlinings, etc. – are encoded using <metamark>. This element is defined in TEI (12.3.4.2) as follows:

By *metamark* we mean marks such as numbers, arrows, crosses, or other symbols introduced by the writer into a document expressly for the purpose of indicating how the text is to be read. Such marks thus constitute a kind of markup of the document, rather than forming part of the text.

<metamark> contains or describes any kind of graphic or written signal within a document the function of which is to determine how it should be read rather than forming part of the actual content of the document. [...]

Unlike marginal notes or other additions to the text, metamarks are used by the writer to indicate a deliberate alteration of the writing itself, such as ‘move this passage over there’. An addition or annotation by contrast would typically concern some property of the passage other than its intended location or status within the text flow. A metamark may contain text, or some other graphic which the encoder wishes to represent, or it may simply consist of arrows, dots, lines etc. which the encoder simply describes.

(<https://www.tei-c.org/release/doc/tei-p5-doc/en/html/PH.html#PH-meta>)

In practice, however, this definition requires slight amplification to account for the marks observed in the *DOFP* documents. The definition highlights that these marks: a) indicate how the text is to be read; b) are “meta” in that they do not form part of the text; c) can be either graphic or textual; and d) may signal a deliberate alteration of the writing. Yet, in the corpus, not all marks indicate how the text is to be read and perform a deliberate alteration of it; on the other hand, certain verbal segments indicate how the text is to be read. When analysing forms of authorial self-communication in the *DOFP*, two axes can be identified, producing four quadrants:

1. Modality: Verbal vs. Non-Verbal
2. Function: Commenting (the author talks about the content or process) vs. Acting (the author performs an operation on the text: moves it, emphasises, re-structures, etc.)

Rather than privileging the function axis, I chose to prioritise modality to avoid ambiguity. Therefore, <note> is used exclusively for verbal expressions and <metamark> exclusively for marks, regardless of function. Consequently, on one hand, marks signalling hesitation, approval, or rejection – although not strictly “deliberate alterations of the writing itself” – are included in <metamark>. On the other hand, words like “backward”, which are instructions for reading the text, were encoded using <note>, as seen before in Example 22.

A mandatory @function attribute and optional @target and @rend attributes are then used to complement the element. Because <metamark> captures the graphic sign itself, it was generally kept separate from the segment to which it refers. In

many instances, a <seg> element was paired with an empty <metamark> specifying @function and @target. Detailed descriptions of each type follow below.

1. Generic linking – These metamarks signal, simply, that a given segment refers to a given location. They may be used in conjunction with <anchor/> elements carrying an @xml:id attribute. Linking arrows differ from replacements (see below). In replacements, a segment is effectively relocated; in generic linking, the author adds textual content and marks where it should be inserted, without moving any pre-existing text.

In the example below, the author first wrote the addition and then drew an arrow pointing to the insertion site.

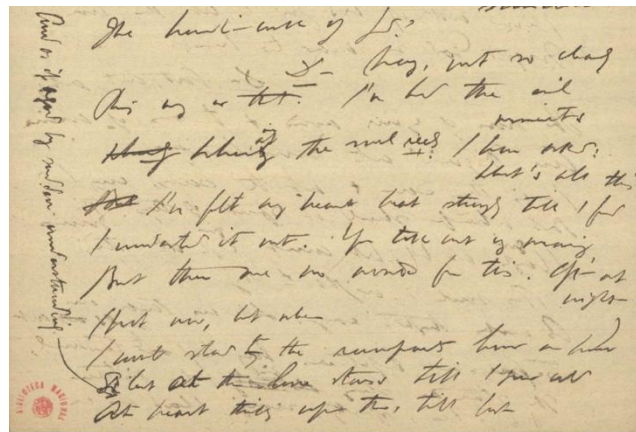


Fig. 8. A linking arrow (BNP/E3, 11^d D-34^r; detail).

```
<lb/>I would stand by the ramparts hour on hour
    <lb/>Silent at the lone stars till I grow old
    <lb/><anchor xml:id="anch_114D-34-1"/>At heart thinking upon this, till last
  </ab>
</sp>
<alt target="#alt_114D-34-3 #alt_114D-34-4 #alt_114D-34-5" mode="excl"/>
</div>
<div type="addition" rend="addColumn bottomText" facs="#_114D-34r-addition">
  <ab>
    <add hand="hnd:FP-blackInkHandwriting" place="margin" rend="writingMode(verticalRL)"><metamark
      function="linking" target="#anch_114D-34-1"/> And as if aged by sudden understanding</add>
    </ab>
  </div>
```

Example 23. A <metamark function="linking"/> element.

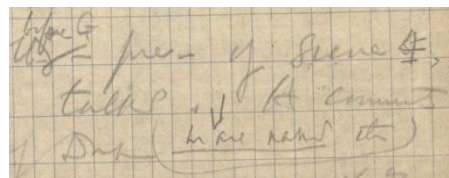


Fig. 9. Another linking arrow (BNP/E3, 11^d D-2^r; detail).

```
talked & <add hand="hnd:FP-graphitePencilHandwriting" place="below"><hi rend="halfCircled"><metamark
  function="linking"/><hi rend="underlined">We are naked</hi> etc</hi></add>
```

Example 24. Another <metamark function="linking"/> element.

In Example 24, the arrow is drawn at the point of insertion before the addition is written. In such cases, @target is not required, since the insertion site is implied.

2. Replacements – Unlike transpositions or alternative readings, which involve the coexistence of multiple textual states and therefore allow TEI to encode both the operation and its result,²⁶ arrow-based replacements do not instantiate a second textual form. In this case, the author does not cancel, rewrite, or reinscribe the segment elsewhere; instead, an arrow merely establishes a reading instruction pointing from an inscribed segment to a physically present locus in the text, which remains blank. This phenomenon is therefore metatextual rather than textual: it organizes the reader's navigation without producing an alternative sequence of characters. Encoding a resulting text or a modification operation (e.g. by means of <mod> elements) would imply a completed textual relocation that is not materially attested and would exceed the evidential limits of a conservative edition. For this reason, the encoding is restricted to marking the affected segment with <seg> and representing the arrow itself as a <metamark> linked to an <anchor/> at the target location, thus preserving the documentary record while leaving any reconstruction of reading order to the visualisation layer (see below).

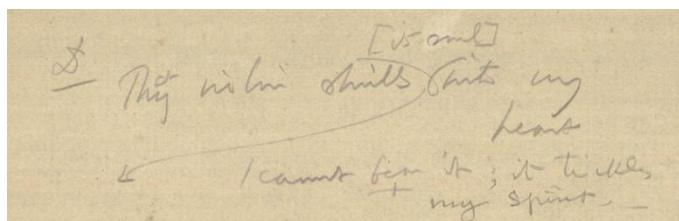


Fig. 10. A replacement arrow (BNP/E3, 11⁵ D-35^r).

```
<div type="textColumn" facs="#_115D-35r-column">
  <sp who="psn:D1">
    <speaker><hi rend="underlined"><persName ref="psn:D1">D</persName></hi></speaker>
    <p>
      <lb>Th<seg type="alt" xml:id="alt_115D-35-1">y</seg> <add hand="hnd:FP-
graphitePencilHandwriting" place="above" xml:id="alt_115D-35-2">at</add> <gap
reason="illegible"><desc>†</desc></gap> shrills <add hand="hnd:FP-graphitePencilHandwriting" place="above"
xml:id="seg_115D-35-1"><hi rend="bracketed">its soul</hi> </add><metamark function="replacement"
target="#seg_115D-35-1 #seg_115D-35-2" rend="arrowBelow"/> <note type="editorial" resp="prj:MJS prj:EL">"its
soul" transposed below.</note> into my
      <lb><hi rend="indented">heart</hi>
      <lb><anchor xml:id="seg_115D-35-2"/> I cannot <seg xml:id="seg_115D-35-3"
rend="hesitationMark">bear</seg> <metamark function="dissatisfaction" target="#seg_115D-35-3"/> it; it tickles
      <lb><hi rend="indented">my spirit;</hi>
    ...
  </sp>
</div>
```

Example 25. Encoding of a replacement (BNP/E3, 11⁵ D-35^r).

²⁶ Transpositions are encoded using <seg> elements, completed by @xml:id attributes, and a <listTranspose> formed by <transpose> and empty <ptr/> elements. In the latter, @target attributes point to the segments to be reordered.

3. Separation marks – lines indicating that a segment is intentionally set apart from the surrounding text are encoded as meta-marks rather than as layout or structural divisions, since their function is not to structure the work formally but to guide reading and interpretation. In the following examples, note the requirement (that will be discussed below) that the element be non-empty to ensure proper processing by EVT:

```
<metamark function="separation" rend="halfRule"> </metamark>
```

Example 26. A `<metamark function="separation">` element.

```
<metamark function="separation" rend="rule"> </metamark>
```

Example 27. Another `<metamark function="separation">` element.

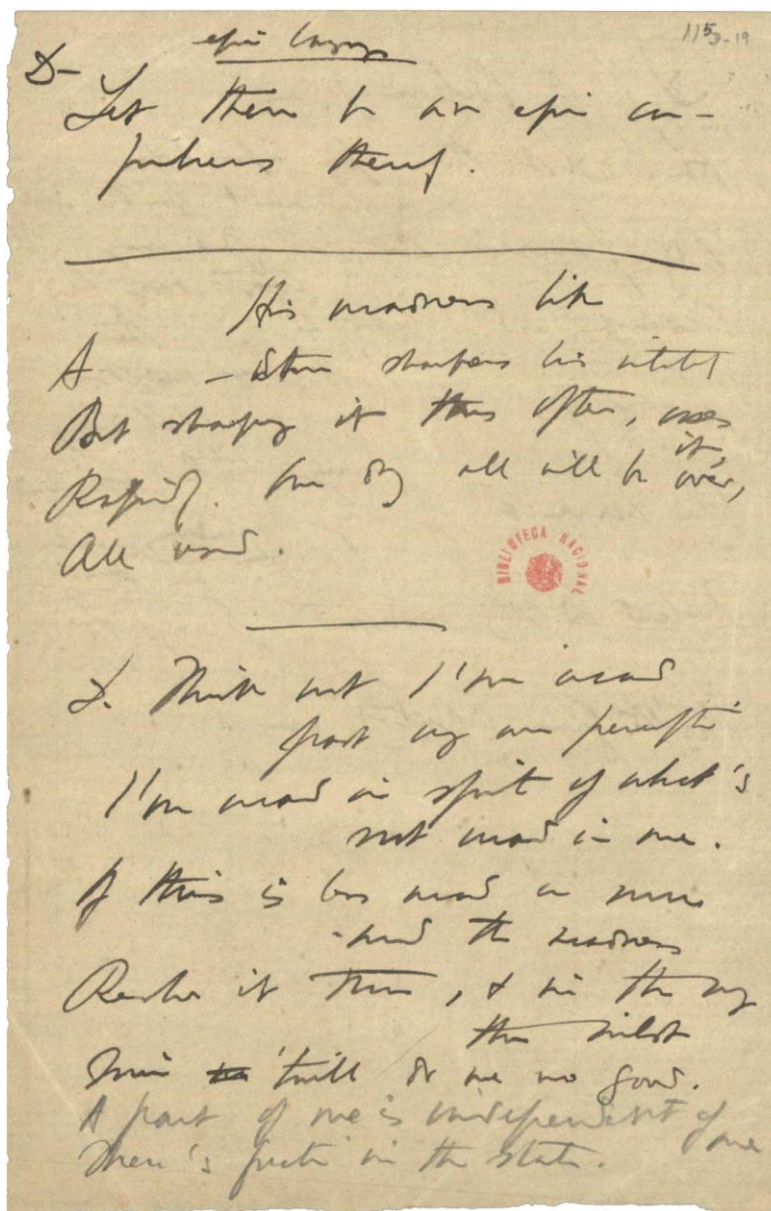


Fig. 11. Separation lines of different widths (BNP/E3, 11⁵ D-19^r).

4. Doubt, approval, and rejection – To express hesitation about certain words or passages, Pessoa typically used strokes, which can be horizontal, when referring to shorter segments, or vertical, when referring to longer passages, crossed by a perpendicular line. Sometimes, he drew question marks. Approval or rejection of passages could be conveyed through symbols such as ticks and crosses.

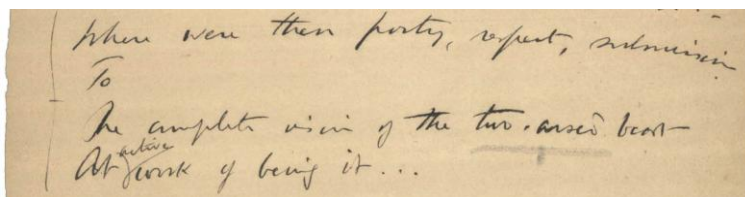


Fig. 12. Hesitation marks, both vertical and horizontal (BNP/E3, 114 D-97; detail).

```
</lb><seg xml:id="seg_114D-97-2" rend="leftVerticalHesitationMark">Where were these poetry, respect, submission
</lb>To <space quantity="45" unit="char"><desc>□</desc></space>
</lb>The complete vision of the <seg xml:id="seg_114D-97-3" hand="hnd:FP-
graphitePencilHandwriting" rend="hesitationMark">two-armed</seg> bear
</lb>At <add hand="hnd:FP-blackInkHandwriting" place="above">active</add> <unclear
cert="medium">work</unclear> of being it...</seg>
</ab>
...
<metamark function="dissatisfaction" target="#seg_114D-97-2"/>
<metamark function="dissatisfaction" target="#seg_114D-97-3"/>
```

Example 28 Use of <metamark function="dissatisfaction"/> in context

In these instances, the text affected by hesitation is encoded within a <seg> element, while an empty <metamark> element – specified by appropriate @function and @target attributes – is placed outside the corresponding structural division.

The possible values of the @function attribute in <metamark> element are summarised in the table below.

Attribute	Possible values	Other possible attributes
@function	linking	@target @rend
	replacement	@target @rend
	separation	@rend
	approval	@target
	dissatisfaction	@target

Table 3. Possible values of the @function attribute in the <metamark> element.

Finally, the encoding process brought to light a further category, comprising marks whose function cannot be determined with sufficient certainty, but which can nevertheless be grouped under a broad notion of graphic highlighting. These include words or phrases enclosed in boxes, circles, or half-circles, as well as brackets, parentheses, underlining, overlining, and similar devices. In such cases, the manu-

scripts provide clear visual salience without offering reliable evidence as to the intention or function of the mark. Encoding these phenomena as meta-marks would therefore entail interpretive commitments that exceed what is materially attested.

For this reason, I considered appropriate to encode such instances with the <hi> element. The function of <hi>, in the TEI Guidelines, is minimal and descriptive: it is defined as a mechanism for encoding highlighting “for reasons concerning which no claim is made” (TEI, 3.3.2). By adopting it in these cases, the encoding preserves the documentary evidence of graphic distinction while avoiding the attribution of intent or function that cannot be securely inferred from the document.

The specific graphic form of the highlighting is captured through the @rend attribute, whose values document the observable visual features of the mark (Table 4).

Attribute	Possible values
@rend	underlined
	underlined overlined
	doubleUnderlined
	circled
	halfCircled
	halfCircledAbove
	halfCircledRight
	boxed
	leftCorneredBelow
	leftCorneredAbove
	bracketed
	parenthetical

Table 4. Possible values of the @rend attribute in the <hi> element.

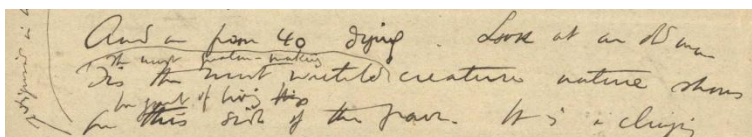


Fig. 13. Half circle above an alternative reading (BNP/E3, 11⁵D-29^v).

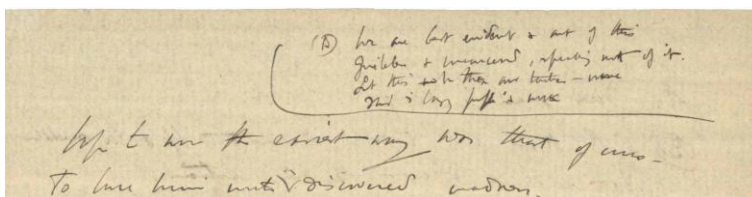


Fig. 14. Segment highlighted with a half-box (BNP/E3, 11⁴D-20^v), providing an example of “leftCorneredBelow” value.

Third iteration: issues arisen from the choice of the interface

The previous sections have focused on the encoding. A further stage of the project concerns the visualisation of the edition, where additional issues emerged. Before examining these issues in detail, a few remarks are necessary regarding the principles that guided the design of the interface. In developing it, in fact, an effort was

made to align each visual feature with a specific conceptual area identified during the modelling of the encoding. While a strict one-to-one correspondence is not always possible, particular care was taken to minimise functional overlap, so that each graphical device would, as far as possible, fulfil a distinct role.

In general, a topographic approach was adopted. Marginal text is positioned in its original location on the leaf in relation to the surrounding main text, rotated text is represented as faithfully as possible. However, a few important exceptions must be noted.

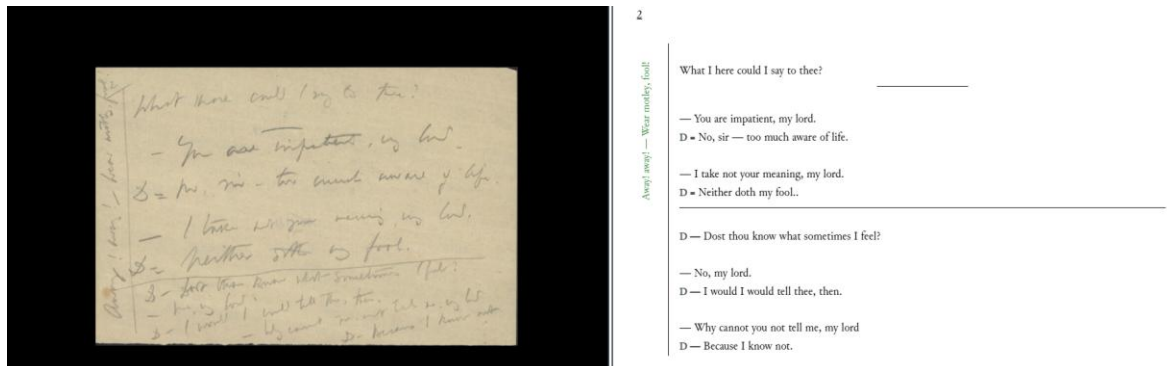


Fig. 15. Topographic approach (BNP/E3, 114D-4v).

First, sections consisting of more than one line rotated 90° to the left cannot be accurately rendered due to limitations of CSS. Consequently, to preserve clarity and coherence, multi-line sections rotated 90° (whether left or right) are displayed horizontally, with a note indicating their original orientation available on hover.

Second, many documents display two or three vertical columns that ideally should be represented topographically. At present, however, such representation would either extend beyond the visible window or reduce the text to a size that compromises legibility and segmentation criteria. In such instances, editorial notes accessible on click are used to clarify the original layout.

Beyond these general topographic principles, other aspects requiring differentiation included:

- Features typical of dramatic texts.
- Onomastic elements.
- Graphic signs present in the original documents (symbols, underlining, arrows, etc.), as well as other elements related to the writing process.
- Editorial interventions.

On the other hand, in any visualisation, the available resources for differentiating features include:

- Symbols.
- Colour.

- Typographical variation (weight, size, inclination, letterform, style).
- Information provided through hover tooltips.

In the *DOFP* model, typographical variation is used primarily to mark theatrical conventions: for example, stage directions are signalled by italics and reduced font size.

Hover tooltips and clickable flagging icons do not interfere with the reading flow or visually overlap with the base text, and are therefore particularly suited to rendering editorial interventions, which must remain distinct from the documentary layer.

The use of specialised symbols associated with genetic editing practices in Portugal was deliberately avoided, in consideration of a broader readership. Two main exceptions stand out: transcription uncertainties (signalled by *) and passages that could not be deciphered (signalled by †).

Both colour and symbols could restrict comprehension to expert readers and create visual noise. Moreover, they were frequently used by the author himself and might therefore overlap conceptually with the documents. For these reasons, they were employed sparingly. Symbols used by Fernando Pessoa are, in general, reproduced, even though certain generalisations proved necessary. While the visual model achieves, hopefully, a clear and coherent reading experience, implementing it was not without challenges.

At the outset of the project, in early 2024, the decision was made to adopt EVT as the visualisation platform. At that time, EVT2 (beta2) was the most recent stable release, though version 3 had already been announced. EVT2 was already installed on the institutional server, and, as far as I know, several ongoing projects at the University of Parma relied on it. From this practical and operational standpoint, its adoption represented a coherent choice, even after version 3 became available later that year.

Initial testing of EVT2, however, revealed a number of technical limitations. For example, the platform does not process empty TEI elements. This posed a problem for elements such as `<space/>`, which is, by definition, empty and is the only TEI element specifically designed to encode spaces intentionally left blank by the author. Its use was therefore called into question. A further limitation concerns the way information is retrieved from the XML files and processed by the platform. In certain aspects, this processing does not fully comply with the TEI Guidelines, making straightforward implementation difficult without substantial modifications of XSLT files. Given the scale of the encoding and the platform's architecture, those modifications to the software itself were neither feasible nor planned within the project timeframe. Targeted workarounds were then implemented to meet the archive's requirements.

These workarounds included suppressing interface elements that were irrelevant to the project's aims, developing a dedicated CSS layer to achieve the desired presentation, and partially restructuring the encoding. Some examples follow.

1. A `<front>` element was added before the `<body>` to contain a discursive summary of the metadata. Although this solution is computationally redundant and not fully compliant with TEI guidelines, it addresses a limitation of the default XSLT, which is not configured to retrieve the `<teiHeader>` of individual files included via `<xi:include/>` – only the `<teiHeader>` of the `dop-main.xml` file.

```
<front>
  <head>Document description</head>
  <p>
    <list>
      <item><hi rend="bold">Repository:</hi> Biblioteca Nacional de Portugal (BNP)</item>
      <item><hi rend="bold">Shelfmark:</hi> Esp. E3, 11-4-D, fols. 21r-21v</item>
      <item><hi rend="bold">Support and material:</hi> Page of laid dun paper, with a medial horizontal fold
and irregular upper, lower and right margins.</item>
      <item><hi rend="bold">Extent:</hi> 1 leaf.</item>
      <item><hi rend="bold">Dimensions:</hi> 222 × 170 mm.</item>
      <item><hi rend="bold">State of preservation:</hi> Well preserved.</item>
      <item><hi rend="bold">Layout (<hi rend="italics">mise en page</hi>):</hi> On fol. 21r, the text is
presented in a single column, opening with a header and consisting of 25 lines. On fol. 21v, the text is arranged in a
single column consisting of 11 lines, with three additional lines written later in the second half of the page.</item>
      <item><hi rend="bold">Hand:</hi> Fernando Pessoa's handwriting in black ink, with interventions in
graphite pencil.</item>
      <item><hi rend="bold">Contents:</hi> Both fol. 21r and fol. 21v contain text related to <hi
rend="italics">The Duke of Parma</hi>, with typewriting interventions in black ink.</item>
      <item><hi rend="bold">Editorial history and encoding:</hi> The transcription was prepared by Nicolás
Barbosa and revised by Jerónimo Pizarro. The analysis of script materials and dimensions was carried out by Carlotta
Defenu. Text encoding in XML-TEI was undertaken by Elena Lombardo, with revisions by Elena Lombardo and
Carlotta Defenu.</item>
      <item><hi rend="small"><hi rend="bold">Last updated:</hi> 2026-01-22 (Elena Lombardo: text
encoding revision).</hi></item>
    </list>
  </p>
</front>
```

Example 29. `<front>` element (BNP/E3, 11⁴ D-21).

The `<front>` was chosen because it is displayed by default through the interface's Info button (Fig. 16), thereby ensuring immediate access to document-level information. While this button was initially intended to be suppressed, it was ultimately retained and repurposed as a functional element for navigation and orientation within the archive.

The `<front>` is encoded using only `<head>`, `<p>`, `<list>` and `<item>` elements, plus `<hi>` elements accounting for some minimal rendering aspects. If the archive were migrated to a different visualization platform in the future, the `<front>` could be removed easily through a simple regex-based find-and-replace operation.

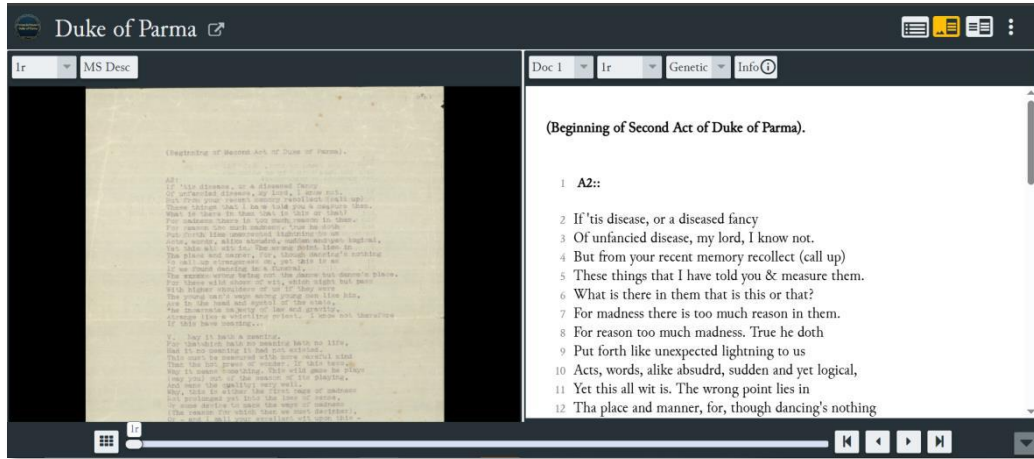


Fig. 16. The interface appearance before the customisation phase

2. As mentioned above, in cases where the author deliberately left blank spaces to indicate the presence of a lacuna to be filled at a later stage, the encoding follows a functional and pragmatic approach. Although TEI provides the empty `<space/>` element to represent blank extents of the material support, EVT2 (beta2) does not process empty elements, which would interfere with both visualisation and interpretation in the reading environment. For this reason, segments that could in principle be encoded as `<space quantity="15" unit="char"/>` were instead completed with a `<desc>` element. This made it possible to insert the editorial symbol conventionally used in such cases (□), which can subsequently be shown or suppressed in the visualisation layer depending on project-specific definitions – while the current implementation does not display this symbol, this behaviour can be modified in future.

3. Because the transcription panel in EVT2 cannot dynamically accommodate extensive marginal commentary – for example, by allowing users to resize the text pane interactively – certain additions that would ideally appear alongside the main text had to be repositioned. In order to preserve legibility, readability was prioritised, and longer additions were relocated when necessary, while maintaining explicit encoding links to their corresponding textual segments and providing a `<note>` to guide the reader through the intended reading flow.

As seen above, `<note>` elements can be both authorial and editorial. Accordingly, in the *DOFP* model, the `<note>` element is used also for editorial annotations. It includes a `@type` attribute (in this case, by default “editorial”), a `@resp` attribute (employing the project’s prefixes to point to the `@xml:id` of `<persName>` elements in *dop-main.xml*), and, where applicable, a `@target` attribute linking the note to the relevant `<seg>` element in the encoded text.

`<note type="editorial" resp="prj:MJS prj:EL">The addition in the left margin is linked to this point by an arrow.</note>`

Example 30. A `<note>` element.

4. Finally, the archive's landing page, containing navigation instructions, was created as a separate XML file and included in the main dop-main.xml container via `<xi:include/>`. Although a full-width page distinct from the standard reading interface might have been preferable, time constraints did not allow for its implementation.

Other limitations remain unresolved: at present, only the container dop-main.xml file is available for automatic download; individual documents cannot be accessed via interface controls and must be instead retrieved manually through from the server via `<https://edizionidigitali.unipr.it/evt/data/text/>`.

Regarding the styling of the transcriptions, as said before:

- Arrows, though encoded in the markup, are not graphically rendered. Their visual reproduction would require dynamic calculation of spatial relationships between distant elements in the layout — functionality that depends on JavaScript-based coordinate processing. Since this edition had to rely exclusively on XML and CSS for rendering, such dynamic graphical connections are not supported in the present implementation.
- `<zone>` and `<surfaces>` elements encoded in the `<facsimile>` were encoded to allow mapping between the facsimile and the transcription, but they do not currently generate a visible effect in the interface.

Ultimately, the *Search* functionality also remains limited: in EVT2 (beta2), the *Create index* button is, to my knowledge, not functional and the search tool only supports anthroponymy and toponymy, while other semantic, structural, or genetic features remain outside its scope.

On creativity in Scholarly Digital Editing

From the outset, the project confirms a long-known assumption: that transcribing is never a purely technical exercise – neither when the aim is to produce a semi-diplomatic digital edition. Choices concerning structure, hierarchy, visualisation, and modes of interaction directly shape how Pessoa's creative process is perceived and accessed. Encoding the unfinished is therefore not merely a technical task, but an editorial practice in which modelling decisions actively participate in meaning-making. In this context, I propose that the archive may be understood as an enriched digital archive – an archive incorporating the interpretative and representational layers of a diplomatic-genetic digital edition.

Within the specific context of SDE, the modelling and encoding of *The Duke of Parma* confirm many of the principles already identified in the field, while simultaneously bringing into focus a dimension that may deserve more explicit theoretical recognition: the centrality of creativity under constraint. Several of the themes

emerged during the project align with well-established features of the digital paradigm in the editorial practice. First of all, the centrality of the modelling phase as the conceptual core of the edition, in which interpretative, structural, editorial and technical decisions converge. Closely connected to this is the inherent iterative nature of the process: modelling confirmed to be neither linear nor definitive, but recursive, requiring constant reassessment as new textual, material, and technical challenges arose. The need to tailor the TEI standard in a few cases further confirmed this dynamic. In particular, the need to adjust the use of the `<rs>/<note>` elements and to create additional attributes exposed an area of genetic dossiers encoding apparently not yet fully addressed in the Guidelines. At the same time, the limits imposed by funding and timeframes decisively shaped the scope of the archive, its degree of refinement, and the level of interpretative ambition that could responsibly be pursued. In this context, the importance of thorough documentation was once again confirmed as an essential guarantee of transparency, possible reusability, and future scalability.

Yet, beyond these familiar aspects, the project has highlighted a relatively underexplored umbrella concept under which they all may be rearticulated: the need for creativity and, to a certain extent, improvisation. When time, funding, and human resources are limited, creativity becomes not an optional enrichment but the only viable methodological strategy. The solutions adopted in the *DofP Digital Archive* – emerging, above all, in the pragmatic adaptations to the constraints of EVT2 (beta2) – exemplify an approach that seeks to extract the maximum analytical value from the tools available. These solutions may not represent the best conceivable configuration in absolute terms; however, they were implemented as the most adequate within the specific material, technical, and institutional context of the project.

In this sense, it may be possible to argue that the archive embodies a *productive* form of provisionality. The unfinished status of the play itself is mirrored in the necessarily open and revisable character of its first scholarly digital edition – not only because future editions, shaped by different interpretative perspectives, will inevitably be proposed, but also because any edition, as Gianfranco Contini famously defined it, is a working hypothesis (CONTINI, 2007: I, 29–30; II, 963). In the analogic paradigm, the editorial work has always been provisional because of the evolving nature of textual scholarship: new witnesses may emerge, new evidence may come to light, and interpretations may shift as knowledge of the work deepens. In this view, editorial provisionality derives from the object itself and from the historically situated condition of scholarship.

While this understanding remains valid, the digital paradigm adds a further layer. Here, provisionality does not depend solely on the possibility of new discoveries or revised interpretations, but also on the technical and institutional frameworks within which the edition is produced. Platforms evolve, tools improve, stand-

ards are refined, and funding structures impose limits that shape what can be implemented at a given moment. As a result, the edition is not only epistemologically provisional, but infrastructurally provisional as well.

It is precisely at this intersection that provisionality may acquire methodological value. Provisionality can become productive when it encourages transparency, reflexivity, and adaptability rather than concealment or rigidity. Taken together, the *DofP Digital Archive* demonstrates that the encoding of fragmentary and unfinished material is not merely once again assessing to which extent an established hierarchical standard such as TEI can be applied to complex textual objects – it is also a process of negotiating between textual complexity and technical affordances. By embracing provisional solutions while maintaining explicit documentation and pursuing structural coherence, the project situates itself within the well-established conception of the SDE as an iterative, evolving construct. The unfinished, rather than being a limitation, thus emerges both as an object of representation and as a methodological condition – one that invites continued refinement, reinterpretation, and collaborative extension.

Bibliography

- BARBOSA, Nicolás (2024). "The Duke of Parma in Fernando Pessoa's Dramatic Corpus: Preliminary Inventory, Classification, and Statistics". *Luso-Brazilian Review*, vol. 61, n.º 2, pp. 56-79. <https://doi.org/10.3368/lbr.61.2.56>
- CELANI, Simone (2024). "The Duke of Parma: Intersections". *Luso-Brazilian Review*, vol. 61, n.º 2, pp. 121-134. <https://doi.org/10.3368/lbr.61.2.121>
- CONTINI, Gianfranco (2007). *Frammenti di filologia romanza: scritti di ecdotica e linguistica (1932-1989)*. Giancarlo Breschi (ed.). Firenze: Edizioni del Galluzzo per la Fondazione Ezio Franceschini. 2 vols.
- DIONÍSIO, João; SOUSA, Maria João (2026). "The Duke of Parma: the organization of the document witnesses". Paper presented at the International Conference Unveiling "The Duke of Parma" (Parma, 23-24 February).
- LIND, Georg Rudolf (1966). "Descobertas no espólio de Fernando Pessoa". *Ocidente*, vol. 70, n.º 334, pp. 57-62.
- LOPES, Teresa Rita ([1977] 2004). *Fernando Pessoa et le drame symboliste. Héritage et Création*. Paris : Éditions de la Différence.
- MARTINES, Enrico (2024). "Presentation of the Special Issue on Fernando Pessoa's *Duke of Parma* and Its Edition Project. *Luso-Brazilian Review*, vol. 61, n.º 2, pp. 1-12. <https://doi.org/10.3368/lbr.61.2.1>
- PIERAZZO, Elena (2019). "Edizione documentaria digitale: rinuncia intellettuale o opportunità scientifica?". *Ecdotica*, n.º 16, pp. 174-185. See the index of the issue: <https://site.unibo.it/ecdotica/en/journal-issues/ecdotica-16-2019>
- PITTELLA, Carlos (2020). "Marino, a tragedy, part I—datable fragments and lists". *Pessoa Plural—A Journal of Fernando Pessoa Studies*, n.º 18, Fall, pp. 593-651. Brown Digital Repository. Brown University Library. <https://doi.org/10.26300/kmse-nv38>
- PIZARRO, Jerónimo (2024). "Editing the Duke: Some Clues and Challenges". *Luso-Brazilian Review*, vol. 61, n.º 2, pp. 135-152. <https://doi.org/10.3368/lbr.61.2.135>
- PIZARRO, Jerónimo; BARBOSA, Nicolás (2026). *Transcribing Fernando Pessoa*. Paper presented at the International Conference Unveiling "The Duke of Parma" (Parma, 23-24 February).
- PORTELA, Manuel; SILVA, António Rito (2017) (eds.). *Arquivo LdoD: Arquivo Digital Colaborativo do Livro do Desassossego*. Centro de Literatura Portuguesa da Universidade de Coimbra, Coimbra. [<https://ldod.uc.pt/>].
- PRICE, Kenneth M. (2009). "Edition, Project, Database, Archive, Thematic Research Collection: What's in a Name?". *Digital Humanities Quarterly*, vol. 3, n.º 3. <https://doi.org/10.63744/jjfxfk55guw>.
- ROBINSON, Peter (2013). "Towards a theory of digital editions". *Variants: The Journal of the European Society for Textual Scholarship*, n.º 10, pp. 105-131.
- TEI Consortium (eds.) (2026) *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Version 4.11.0. [<http://www.tei-c.org/Guidelines/P5/>]. Last updated on 18th February 2026, revision 358d2e48e.
- TERLINDEN, Anne (1978). "Fernando Pessoa e a sua obra em inglês". *Persona*, n.º 2, pp. 65-68.
- ZENITH, Richard (2021). *Pessoa. A Biography*. New York: Liveright Publishing.

ELENA LOMBARDO is an Invited Assistant Professor at the School of Arts and Humanities of the University of Lisbon. She earned a degree in Linguistic and Cultural Mediation from the University of Padua (2011), an MA in Portuguese Language and Philology from the University of São Paulo (2015), and a PhD in Textual Criticism from the University of Lisbon (2025). Her doctoral thesis focused on the study and scholarly digital edition of a Portuguese chronicle related to the reign of King Sebastian. She is a member of the Centre of Linguistics of the University of Lisbon (Philology Group) and coordinates the project *Sebástica Manuscrita*, as well as the initiative *From Parchment to Computer: Editing Manuscripts in the Digital Age*. Her work bridges traditional philology and digital approaches and her research interests include textual genesis and transmission, manuscript culture, and the development of scholarly digital editions.

ELENA LOMBARDO é Professora Auxiliar Convidada na Faculdade de Letras da Universidade de Lisboa. Licenciou-se em Mediação Linguística e Cultural pela Universidade de Pádua (2011), obteve o grau de Mestre em Filologia e Língua Portuguesa pela Universidade de São Paulo (2015) e doutorou-se em Crítica Textual pela Universidade de Lisboa (2025). A sua tese de doutoramento consistiu no estudo e na edição filológica digital de uma crónica portuguesa sobre o reinado de D. Sebastião. É membro integrado do Centro de Linguística da Universidade de Lisboa (Grupo de Filologia) e coordena o projeto *Sebástica Manuscrita*, bem como a iniciativa *Do pergaminho ao computador: editar manuscritos na Era Digital*. Desenvolve investigação que articula a filologia tradicional e as abordagens digitais, com especial enfoque na génese e transmissão dos textos, na cultura manuscrita e nas edições filológicas digitais.