Digital Encoding as a Hermeneutic and Semiotic Act: The Case of Valerio Magrelli

Domenico Fiormonte <fiormont_at_uniroma3_dot_it>, Università Roma Tre, Dipartimento di Italianistica
Valentina Martiradonna <valem1078_at_hotmail_dot_com>, Università di Roma, La Sapienza
Desmond Schmidt <desmond_dot_schmidt_at_qut_dot_edu_dot_au>, Queensland University of Technology, Information Security Institute

Abstract

In this article we propose different methods of encoding, according to the TEI Guidelines, three different cases of genetic or compositional textual variants found in the autographs of the Italian contemporary poet Valerio Magrelli. These encoding experiments reflect the diverse nature of the artifacts and represent a critical assessment of the effectiveness of present encoding practices for the multidimensional and pragmatic aspects of authorial drafts. Thus far, it seems that the TEI has yet to offer a convincing theoretical model and adequate practical solutions for representing the complex temporal structures normally present in manuscripts, and in fluid textual traditions in general. Our conclusion is that there is a potential conflict between the linear and hierchical nature of current formal language systems such as XML, and the intrinsic dynamic nature of the writing process. In such cases we may have to rethink present models of document modeling, and to develop, within an adequate epistemological framework, a new theory of digital text.

1. Fluid Text and Markup Languages

In this paper we shall try to uncover, through some experiments in digital text-encoding, the complex multidimensional and interactive reality of the process of composition. This complexity emerges as a result of problems encountered in the description and representation of a series of autographic literary texts. The historical and theoretical background of our work is represented by various philological-critical schools such as the Italian criticism of variants (*critica delle varianti*), French genetic criticism and the Anglo-American tradition of textual bibliography.[1] Each one of these traditions have developed, since the beginning of the 20th century, different critical instruments and approaches; however they all seem to share a common idea: a literary work (and in general every "script act") can be regarded not merely as a product, but as the result of a dynamic process of interaction between several factors and influences of a linguistic, cultural and social nature.[2] This attention to processes is also studied in pragmatics [Austin 1976], whose contribution to the study of textual fluidity has been explicitly recognised in the text-critical domain in a recent article by Peter Shillingsburg [Shillingsburg 2006a].[3] Although there is no room here to investigate the pragmatic aspects of autographic writing, it is obvious that the typical illocutory and dialogic nature of the textual and graphic-semantic elements of the *avant-texte* constitutes a relatively fertile area for pragmatic analysis.[4] The *avant-texte* is a *communicative system*: the signs on the manuscript page are primarily interpretations of "actions," and the dialogue involves the author not only as a subject in multiple roles (author/reader, author/corrector, author/critic, etc.), but also in triangular relations such as editor/author/editor as in the case of proof-editing.

One of the main issues that emerged from our encoding experiment can also be formulated as a question: can the multidimensional and pragmatic nature of different writing stages/sketches be represented with the help of digital instruments? As we will see, a tentative answer to this question leads to the admission that any transformation from paper to bits, apart from leading to certain developments, all potentially "intrusive" for the document, is no ordinary

1

2

event, nor is it technically neutral. The first stage is the conversion of the textual content and structure of the original document into digital form. This process, far from representing a simple copying of a document from one medium to another, is actually a hermeneutic and semiotic process, for in the moment in which a text is transcribed through the selection and the use of markup this creates meaning in itself. It is also a pragmatic process, since markup is not only able to represent but also to *create* actions [Renear 2005, 28]. In the words of one of the Italian pioneers of humanities computing:

> We must understand and keep in mind that at the moment of data entry the text (the entire quantity of information contained within it) is being entrusted to a different channel from that in which it has survived until now (if one of the elements of the system of communication changes, the others necessarily change as well). We must also always remember that transcription is in any case an interpretative act that manifests itself at the very moment in which we take any decision: from the simplest — is this element a full stop, the end of an abbreviation or a dash? – to the more complex — is this other case a verse, a single verse on two printed lines, or two verses? [Gigliozzi 1998, 228]

Such "decisions" imply a perspective, a selection of aspects, a method of analysis and a choice of an encoding model, to arrive at a representation of the text out of so many possibilities — an encoding that we realize through a markup language permits us to formally describe the structure of a text, and to analyze the data in depth. Its utility will be in proportion to how much information it can set out, include and preserve. However, it will not be possible, and would hardly be desirable, to represent all aspects an ideal reader would see in a text. The printing bias has forced us to think of a text as a stable product, but if we either look at the different historical representations of a given text or at its documented writing stages, we realize that there is not one text, but many texts, as many as there are mechanisms of writing, material production, intertextual paths and methodologies of reconstruction. [3]

Nowadays the most powerful device for the representation and digital analysis of literary texts is XML (Extensible Markup Language), which provides a rigorous syntax for representing the deep structure of a text. XML is a metalanguage, that is to say a system for defining tags that describe the role played by every element within a text. Apart from technical advantages (portability, standardisation, flexibility, representational power) markup languages have been credited with allowing the scholar to make his own interpretation of the text more explicit [Ciotti 2005, 9–42]. [4]

Of course, markup languages, like any other technological instrument, are not perfect. They impose a hierarchical structure, which is normally used as a container for the text.[5] This can lead to major difficulties in cases where the text cannot be reduced to such a structure. The "well-formedness" of XML, the requirement that every element, except for the document root, must be contained in some other element, gives the text an explicit and unambiguous structure. This principle is in contrast with what Buzzetti defines as the "dynamic instability of literary texts" [Buzzetti 2002]. A number of recent papers have underlined the seriousness of the problems that such dynamic instability poses for XML markup. For example, [Vetter and McDonald 2003] explore the main techniques for recording variants taken from the TEI Guidelines, as well as several ad hoc methods for interlinking them in the poetry of Emily Dickinson. They conclude rather negatively that no method for recording Dickinson's variants proves to be "entirely satisfactory" [Vetter and McDonald 2003, 151]. [Vanhoutte 2006] likewise, in applying modern encoding techniques to record the variation of the manuscript antecedents of a classic Flemish novel, remarks, "The modern manuscript shows a much more complicated web of interwoven and overlapping relationships of elements and structures" [than the printed book]. And although not a case of a modern manuscript, [Bart 2006] describes the problems of encoding variation in the Ht manuscript of Piers Plowman, such as variation in line numbering and the encoding of variants between physical versions, for which the author had often to resort to "the shameless application of technical duct tape." [5]

Specific cases can always be dismissed as anomalies, but these problems appear to be serious enough to warrant further investigation. Since we work on modern manuscripts, and on visualising the encoding of such texts, we are interested in discovering whether the *Text Encoding Initiative Guidelines* (TEI) are in fact sufficient to encode modern manuscript material. Investigation of specific cases does have the advantage of providing a wealth of fine detail that can't be derived from purely theoretical or technical investigations. [6]

The TEI offers a rich collection of guidelines, originally for the encoding of ancient texts, although in later editions it has included specific guidelines for the textual phenomena of modern and contemporary texts [Burnard & Bauman 2007, Ch. 11]. The encoding of these elements, which Vanhoutte calls the "temporal unity" of writing, is often problematic [Vanhoutte 2003, 12] or "very complicated, if not impossible" [Pierazzo 2007, 151]. Any movement of the text *forwards* or *backwards* in time (cf. below § 2.3) naturally tends to generate overlapping structures. However, the spatial dimension can also represent a problem: we need only think of modern poetic texts, in which it is often very difficult to recover the structure of a composition in strophes and verses. The following paragraphs are an attempt to confront these clashes of theory and practice via a case study: the encoding of three compositions and their various writing phases by the Italian poet and writer Valerio Magrelli.

7

2. TEI-XML Encoding of the Poetry of Valerio Magrelli

Our encoding experiment focuses on the three poems ("Molto sottrae il sonno alla vita," "Essere matita è segreta ambizione," "Ecco la lunga palpebra della donna") written between 1975 and 1979, which then appeared in the first section ("Rima palpebralis") of the anthology *Ora serrate retinae* edited by Feltrinelli in 1980.

8

For each poem we have been able to access and scan the author's original version, which in its natural state *in fieri* poses the most important and significant problems of representation. There are also further typewritten versions, gathered and *recatalogued* by the author in various notebooks,[6] in which we find again the strong presence of textual variation, the definitive printed edition and in some cases a printed version in French. The author continued to edit his own texts until the appearance of the printed version, but our work is concentrated on the first two stages of the process of writing the three poems mentioned above: the two autographs of the first two poems ("Molto sottrae il sonno alla vita" and "Essere matita è segreta ambizione") and two typescripts with corrections in the author's hand in the case of the third ("Ecco la lunga palpebra della donna").[7]

9

2.1 Space and Time. Encoding of the Autograph of "Molto sottrae il sonno alla vita"

In our encoding we have tried to represent the movement of writing, whose intention was to produce, by each new editorial action and variation, a new text and a new context. As Allen Renear noted [Renear 2005], it is possible to apply the categories of pragmatic analysis, and in particular the notion of illucotory acts, to better analyse and define the scope and specific uses of markup. As will be seen, we are forced to use markup to favour the illocutory force of the author's own text (its dialogic-contextual dimension) rather than limit ourselves to recording phenomena in a sequential and, as it were, "external" fashion. In philological terms this means that after a process, here outlined, consisting of sketches, proofs and discards, we have decided to favour the diachronic aspect, establishing different phases of writing and rewriting, and emphasising the significance of the author's interaction with the text. As a result, since the choice of encoding is never neutral, either the graphical aspect of the page or the synchronic actions of writing will be sacrificed. Also because, as we noted above, as a result of the limits imposed by the instruments of encoding, often the simultaneous representation of both aspects becomes excessively difficult if not absolutely impossible.

10

The *TEI Guidelines,* version P5,[8] produced by [Burnard & Bauman 2007], is available in the form of several customised modules such as Drama and MS. These contain a selection of various parts of the overall Guidelines and help reduce the complexity and size of the tagset. The TEI also provides specific elements for the representation of poetic texts. The basic plan is the following:

11

```
<lg>
        <l>verse</l>
        <l>verse</l>
        <l>verse</l>
</lg>
```

12

The `<lg>` tag groups the strophe as a unit, while the `<l>` element specifies the verse.[9] In our analysis we chose not to

13

use this model (which is more applicable to the representation of non-contemporary poetry), primarily because the author's originals in this case don't always contain an explicit division into verses, which was often added by the author at a later time or not at all; and secondly because this subdivision is itself subject to alteration and hence must be considered as part of the phenomenon of textual variation.

We see here the first example of the complexity of the textual phenomena from the encoding of a fragment of the autograph of "Molto Sottrae" Figure 01, the poem that opens the collection both in the 1980 edition and in that of 1996 [Magrelli 1996, 7]:

2.1.1 First Encoding

The first attempt at encoding had the objective of describing the physical structure of the page, and considered the autograph at one instant in its history rather than as the result of a succession of events. It was immediately noticed (fortunately for us) that the original layers of correction could be identified by the use of two colors: a red pen for the first draft ("Il sonno è l'indiscreto ospite // irresistibile") and a blue pen for corrections and additions ("<E> si allarga <nel corpo> come un secondo corpo intollerabile..."). As we will see later, thanks to the use of color we can discern with certainty at least four successive stages of corrections by the author, assuming they were carried out at different times.

Since we regarded the rigid division into line-groups (`<lg>`) and verses (`<l>`) as inadequate, we initially marked the end of each "typographic" line with the empty element `<lb/>` (line break) and, since it was impossible to consider this as a suitable mark for a strophic unit, we enclosed a group of verses or the text of the autograph in general within a generic element `<ab>` (anonymous block). In addition, we also marked the end of the poetic verse with another empty element `<milestone/>`, qualified by an attribute `unit="verse"`. Here is the result of the first attempt:

```
<handNote xml:id="blue" scribe="Magrelli" medium="blue biro"/>
<add hand="#blue" place="left">E</add> Il sonno
<subst>
   <del hand="#blue" type="overstrike">è l'indiscreto
ospite<lb/>irresistibile</del>
   <add hand="#blue">si allarga <add hand="#blue">nel corpo/<milestone
unit="verse"/>
      come un secondo<lb/> corpo intollerabile./<milestone unit="verse"/>
   </add>
</subst>
```

We resorted to the empty elements `<lb/>` and `<milestone/>`, representing respectively the end of the typographic line and the end of the verse, to avoid problems of overlap.[10] In the autograph, in fact, the portion of text cancelled by the stroke of a blue pen ("è l'indiscreto ospite // irresistibile") crosses the end of the typographic line and, if we had used a tag like `<l>verse</l>` or even the more generic `<seg>verse</seg>` to represent the verse unit, we would have generated a case of overlap: cancellation followed by replacement in this case involves a unit longer than a verse.

In any case, although the differentiation between the end of the verse and the end of the typographic line by means of the empty elements `<milestone unit="verse"/>` and `<lb/>` resolves the problem of overlap, it also generates a redundancy. These elements in fact fulfil the same function: both mark the end of a portion of text without qualifying it as belonging to either. This poses a problem whenever the two elements refer at the same moment to a common piece of

text.

Even though it is possible to represent the author's own substitutions, deletions and insertions via the tags `<add>` (addition) and `<del>` (deletion), this choice of markup cannot represent the chronological sequence of corrections. For example, it doesn't take into account the fact that the initial verse extends onto two lines (although it is one unit), or the subdivision of the initial verse into two, after *"nel corpo/"* was added, or the subsequent introduction of a metrical structure (the subdivision into verses indicated by Magrelli's use of the forward slash /). In the proposed encoding the state of the metrical scheme was regarded as already decided: the insertion of the metrical structure is anticipated and reshaped by the editor, when in reality it happens at a later point in the process. Also, by this selection we lose important information concerning the relations between the author's interventions and their order. In conclusion, it is not possible to represent both phenomena: the metrical (spatial) structure and that of the (temporal) variation, without violating the syntax of XML. And in the case of the other canto it would perhaps not be appropriate to speak of the typographical design of the page for an object like a manuscript, which is not strictly speaking a publication.

20

The autograph is in fact considered as an open semiotic system, a "field of action" [Ferrer 2002, 52] interdependent on its various possible concrete realizations (among which are printed publications, whether intermediate or definitive).

21

2.1.2 Second Encoding

In this second attempt we will utilise the elements defined for the transcription of the critical apparatus of a manuscript, merging the elements defined in chapters 11 and 12 of the TEI Guidelines.[11] An apparatus is usually understood as an instrument that records the sequence of variants between multiple texts, whereas here it is used to represent the variants within a single manuscript. In our case then we have represented textual variation with the `<app>` (apparatus) tag which, instead of recording, as in its standard use, variants from several manuscripts, here contains the different stages of composition of a single text (our fragment). Each phrase is represented by the tag `<rdg>` (reading), qualified with a `varseq` attribute (variant sequence), which in its turn supplies a number to indicate the sequence. After having identified the various levels of stratification in the text, which occur uniformly throughout the autograph, we assigned a successive number to each reading via the attribute `varseq`. The number indicates the stage of intervention, according to our interpretation, to which the segment of text contained there belongs. As will be seen, in keeping with our intentions, we will represent both the stages of the writing process and the interventions by the author on the metrical structure of the poetry:

22

```
<app>
  <rdg varSeq="1">Il sonno è l'indiscreto ospite<lb/> irresistibile.</rdg>
  <rdg varSeq="3"><add hand="#blue" place="left">E</add> Il sonno
    <del hand="#blue" type="overstrike">è l'indiscreto ospite<lb/>
irresistibile.</del>
    <add hand="#blue" place="top">si allarga come un secondo<lb/>
       corpo intollerabile</add></rdg>
  <rdg varSeq="4"><seg type="l">E il sonno si allarga
    <add hand="#blue" place="supralinear">nel corpo</add></seg>
    <seg type="l">come un secondo<lb/> corpo intollerabile</seg></rdg>
</app><lb/>
```

We have chosen to mark the verse by the `<seg>` element, which is used generically to mark a segment of uncategorised text, giving it an attribute `"l"`, which identifies that segment as a verse unit that we have reconstructed. Examination of the whole autograph reveals four stages: the red biro represents the first form given to the composition by its author, and what may be the second stage consists of the author's interventions — deletions, rewritings and corrections — with the same implement. The use of the blue pen indicates a later stage, the third, followed by corrections on top with the same instrument (fourth stage). In the chosen fragment Figure 01, however, only three stages are visible: the first, the third and the fourth.

23

The initial verse is thus the following:

24

1. *Il sonno è l'indiscreto ospite irresistibile* (red biro).

Subsequently the author inserted with the blue biro an *E* and modified the verse, deleting a portion of it (*è l'indiscreto ospite // irresistibile*) and added some new text (*si allarga come un secondo corpo intollerabile*). Hence we have

3. *E il sonno si allarga come un secondo // corpo intollerabile*

The result of the final revision is the integration of *"nel corpo/"* and the introduction of a metrical structure (the character */* indicates the end of a verse):

4. *E il sonno si allarga nel corpo / come un secondo // corpo intollerabile /*

In the encoding, we have chosen to represent not only the actions of the author and hence the physical data (cancellation, rewriting, insertion of elements which subdivide the verse unit) but also the consequences of these actions (the first portion of text is substituted by another, forming a new verse; this is turn undergoes variation to form two new verses). Then, by eliminating the empty milestone element and inserting the `<seg type="l">` tag, we make one reconstruction of the verse explicit, while marking the end of the typographic line with `<lb/>`.

In this way, we resolve some of the problems that emerged from the first encoding:

<div style="text-align:right">25</div>

1. We represent textual variation with the `<app>` tag and indicate the sequence of the author's corrections with the `varseq` attribute;
2. We contain the textual variation within the element `<seg type="l">`, representing the reconstruction of the verse unit;
3. We avoid the redundancy of empty elements by specifying which unit undergoes textual variation.

It is obvious that these choices not only constitute a specific (and questionable) interpretation of the process of writing, but they do not account for the external aspect of the autograph. And from the point of view of a palaeographer or an archivist, this would be a serious loss of information. Insofar as our proposed encoding can be revised, perfected and extended, we find ourselves facing the central theoretical crux of the representation of digital documents: encoding has the virtue of requiring us to explain and justify our choices, but at the same time this assumption of responsibility (useful and correct in the eyes of the scientific community) declares, as it were, its epistemic limit: not all knowledge can be represented with the present digital tools.

<div style="text-align:right">26</div>

2.2 Alternative or Variant? Encoding of *Ecco la lunga palpebra della donna*

The second composition we examine allows us to observe which phenomena shed light on the encoding of deletions, substitutions, and insertions within the same autograph, in particular the problematic distinction between hierarchies of variants.

<div style="text-align:right">27</div>

As in the preceding case, it is possible to reconstruct the chronology of Magrelli's interventions by examining the graphical elements, and see how a change in pen color indicates a subsequent addition, or by examining the spatial elements, such as the place where the author puts the added text. But such interpretations are insufficient to interpret and encode the process of composition. In the autograph of *Ecco la lunga palpebra della donna*, we are faced with words that are cancelled and substituted not by another word but by two different ones, by additions above the line (supralinear) or below the line (sublinear). In the fragment of the poem that we chose to examine Figure 02 the text is spread over five lines (which we shall only call verses by the laws of post hoc), and is rich in intratextual and pragmatic phenomena of both the textual and graphical kinds (e.g. arrows, marks of emphasis, cross references etc.) In particular the adjectives were subject to close attention by the author and could be reduced to two lists of variants, which Magrelli drew up within the autograph (of which one is visible in Figure 02: the four adjectives written in the bottom right in capitals).

<div style="text-align:right">28</div>
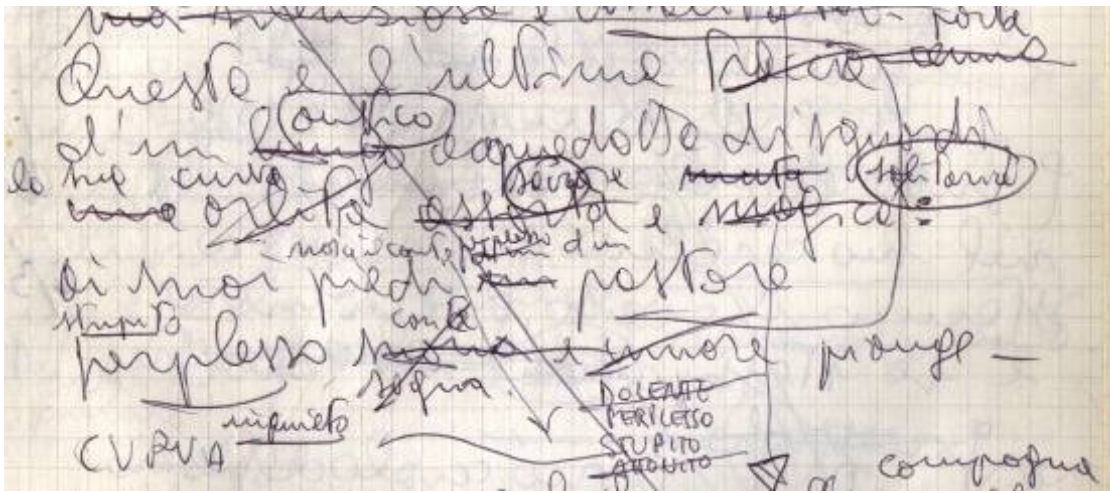
As can be seen, the adjectives *stupito, perplesso* and *inquieto* on the last line (bottom left of Figure 02) could have been added at the same moment, or *stupito* and *inquieto* could have been inserted subsequently and could be considered as the two variants parallel in time to *perplesso.* For this reason we have considered them as a list of variants within the line, whose chronological succession we will not attempt to reconstruct. Immediately following this are two other variants, *canta* and *sogna,* which serve as replacements for *suona.* In this case one can suppose a chronological sequence (*suona* replaced by *canta* and *sogna*), which explicitly retains the value of the coexisting variant alternatives. Here is the encoding of the entire fragment:

```
1       <handNote xml:id="M" scribe="Magrelli"/>
        <!-- ... -->
                <rdg varSeq="2"><subst><del hand="#M"
                type="overstrike">traccia</del><add hand="#M"
5               place="right">cenno</add></subst></rdg>
                <rdg varSeq="3"><subst><del hand="#M"
                type="overstrike">cenno</del><add hand="#M"
                place="supralinear">porta</add></subst></rdg>
        </app>

10      </seg><lb/>
        <seg type="l">d'un
                <app>
                        <rdg varSeq="1">lungo</rdg>
                        <rdg varSeq="2"><subst><del hand="#M"
15                      type="overstrike">lungo</del><add hand="#M"
                        place="supralinear"><emph rend="circle"
                        n="adj_2">antico</emph></add></subst></rdg>
                </app> acquedotto di sguardi,
        </seg><lb/>
20      <seg type="l">
                <app>
                        <rdg varSeq="1">una orbita assorta e magica</rdg>
                        <rdg varSeq="2">una<subst><del hand="#M"
                        type="overstrike">orbita</del><add hand="#M"
25                      place="supralinear">curva</add></subst> <del
                        hand="#M" type="overstrike">assorta</del> e <del
                        hand="#M" type="overstrike">magica</del> e <add
                        hand="#M" place="supralinear">muta</add></rdg>
```

```
                    <rdg varSeq="3"><del hand="#M
30                  type="overstrike">una</del><add hand="#M"
                    place="supralinear">la sua</add>curva <emph
                    rend="circle" n="adj_3">sacra</emph> e
                    <del hand="#M" type="overstrike">muta</del>
                    <add hand="#M" place="supralinear"><emph
rend="circle"
35                  n="adj_4">solitaria</emph></add>:
                    </rdg>
                </app>
        </seg><lb/>
        <seg type="l">ai suoi piedi
40          <app>
                    <rdg varSeq="1">un pastore</rdg>
                    <rdg varSeq="2"><add hand="#M"
                    place="supralinear">nasce il canto
                    <note type="arrow" resp="#M" place="foot">
45                  <list type="simple" n="adj_list1">
                            <item><del hand="#M"
                                type="overstrike">DOLENTE</del>
</item>
                            <item>PERPLESSO</item>
                            <item>STUPITO</item>
50                          <item><del hand="#M"
                                type="overstrike">ATTONITO</del>
</item>
                    </list>
                    </note>
                    d'un</add> pastore</rdg>
55          <rdg varSeq="3">nasce il canto
                    <note type="arrow" resp="#M" place="foot">
                    <list type="simple" n="adj_list2">
                            <item><del hand="#M"
                                type="overstrike">DOLENTE</del>
</item>
60                          <item>PERPLESSO</item>
                            <item>STUPITO</item>
                            <item><del hand="#M"
                                type="overstrike">ATTONITO</del>
</item>
                    </list>
65                  </note>
                    <del hand="#M" type="overstrike">d'un</del>
                    <add hand="#M" place="supralinear">perplesso
                    d'un</add> pastore</rdg>
                </app>
70          </seg><lb/>
            <seg type="l">
                    <list type="varianti" resp="#M">
                            <item><emph rend="underlined"
```

```xml
                                    n="adj_5">perplesso</emph></item>
75                      <item><emph rend="underlined"
                                    n="adj_6">stupito</emph></item>
                        <item><emph rend="underlined"
                                    n="adj_7">inquieto</emph></item>
                    </list>
80                  <delSpan hand="#M" type="overstrike"
    spanTo="#delend01"/>
                    <app>
                            <rdg varSeq="1">suona e muore</rdg>
                            <rdg varSeq="2"><del hand="#M"
    type="overstrike"
                                    n="d2">suona e</del>
85                              <add hand="#M">
                                <list type="varianti" resp="#M">
                                        <item>canta</item>
                                        <item>sogna</item>
                                </add></list><anchor
    xml:id="delend01"/>
90                              <del hand="#M"
    type="overstrike">muore</del><add

                                hand="#M" place="right">piange.
    </add>
                            </rdg>
                    </app>
                </seg><lb/>
```

Here also we have used the TEI Guidelines for transcription of an apparatus criticus, enclosing the textual variation in the `<app>` tag, and assigning to each reading (`<rdg>`) of the verse a sequence number.[12] In the final verse (lines 71-94) we find the alternation of two variants, *canta e sogna* (lines 87-88), which replace the cancelled word *suona;* in the final passage we use the empty element `<delSpan/>` [13], which marks a portion of cancelled text, to indicate the deletion via an oblique line of the three variants altogether (*suona, canta* and *sogna*). This triple cancellation is anchored via the `<anchor/>` element to the point at which the cancellation finishes (line 89).

[31]

To represent the references between the list and the elements in the text we assigned to each adjective an identifier using the n-attribute (preceded by the "adj" attribute: `adj_1, adj_2,` etc.) The adjectives recur several times in the text and the author attached a particular importance to them, circling or underlining them. To render this emphasis we have used the `<emph>` tag, giving it a `rend` attribute which explains what type it is (circled or underlined). The encoding of the fourth verse (lines 39-70) presents one solution to a complex set of interrelationships among elements in various areas of the autograph. Subsequently, probably during the insertion of *nasce il canto,* the author inserted an arrow, which refers to a list of four adjectives written in capitals at the foot of the page (Figure 02, bottom left). This type of indication can be interpreted as an "action" of the author directed towards producing certain effects — but not results, because the author is, in a certain sense, also engaged in a dialog with himself. To record this dialectic we have chosen to represent the arrow with a note, via the metatextual `<note>` element (line 44), giving it the attribute `type="arrow"`, and inserting it in the text containing a list of the adjectives as alternatives (each distinguished from the others by `<item>`). Since we maintain that the list/note of four adjectives was inserted during the second draft, we assign it to the second version (`<rdg varSeq="2">`). In the facsimile the arrow is joined to the o of *[nasce il] canto* (added above the fourth line of the autograph). In summary, our reading of the fourth verse treats the variants within the text as components (`<item>`s) of a list (`<list>`), which we connect back to a collection of alternatives placed by the author at the foot of the page. Obviously, this solution, which thus inverts the spatial distribution of the elements in relation to how
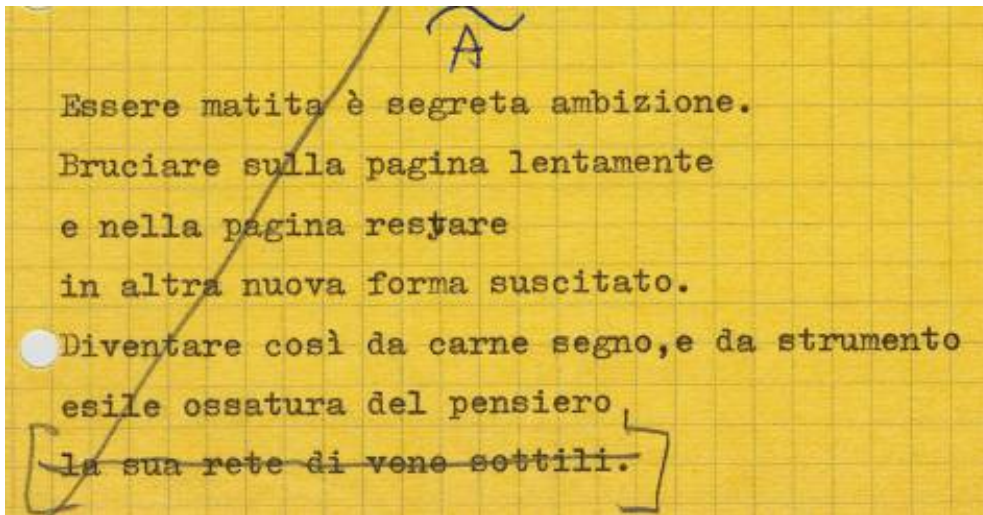
[32]

they are presented in the autograph (the list occurs at the end of the fifth verse and not within the fourth as in the encoding), is one possible workaround for representing such complex phenomena. This is a further example of how encoding of the writing process can imply, through the remodelling/reconstruction brought about by markup languages, the overturning of the spatial dimension. In other words, what appears in the material reality of the document as a particular region, determined visually, in the pragmatic dimension of the process may be located on several temporal levels. Since to encode means to select one point of view, in our case we are forced to represent, through the linear instruments of markup, something which is not linear by its very nature. This is a dimension which the author indicates "performatively" through signals that are not exclusively textual, as if to show that the dimension of the process cannot be cognitively reduced to a hierachical sequence.

2.3 Backwards in Time: Encoding of *Essere matita è segreta ambizione*

It is appropriate that this last example of encoding explores the relationship between writing and time. In this case we shall analyse a composition in which the material succession of witnesses does not coincide with the temporal succession of the various writing phases. We possess numerous versions of *Essere matita è segreta ambizione,* which reveal, in addition to a complex process of composition, a classification peculiar to Valerio Magrelli. Each witness is contained in a notebook representing a specific phase of writing. Altogether there are eight witnesses: the autograph, the second typewritten version, contained in the notebook "Foglietti I," the photocopy of the second version (also in "Foglietti I"), the third in "Foglietti II," the fourth in "Libro — parte I," the fifth printed, the sixth, which will be the definitive printed version, and a printed version in French. What follows are the images of two versions of the first part of the poem *Essere matita è segreta ambizione:*

1. Second version (which we call A):

2. Sixth version (which we call B):

A

Essere matita è segreta ambizione.

Bruciare sulla pagina lentamente

e nella pagina restare

in altra nuova forma suscitato.

Diventare così da carne segno,e da strumento

esile ossatura del pensiero

la sua rete di vene sottili.
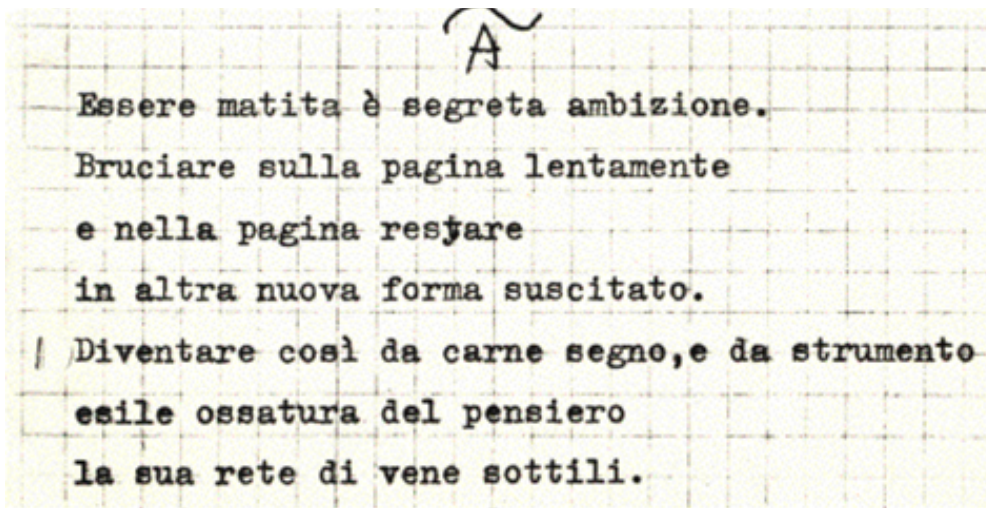
The sixth version (B) is a photocopy of the first version (A). It thus follows it in time, but, as can be seen, it also precedes the insertion of later corrections in the original. The cancellation in pencil of *la sua rete di vene sottili* in fact will not appear in subsequent versions, not even the final one [Magrelli 1996, 13]. The temporal inversion then generates another case of overlap: the correction in A before B is a typical case of "genetic order" [Ferrer 1995, 143], which contrasts with the material order. In reality we are faced with three dimensions of the text:

1. The yellow page preceding the corrections (A0)
2. The photocopy without corrections (B)
3. The yellow page with corrections (A).

In the encoding at this point we have two options, both probably unsatisfactory: to emphasize the genetic aspect, representing the writing phases from A0 (no longer in material existence, but genetically real), or to catalog the witnesses, maintaining their material succession without taking account of the writing phases. We chose the second of these two options for our proposed encoding, although we also decided to highlight the temporal sequence by inserting an editorial note at the point where the correction occurs and so establish a link with the second photocopied version, explaining that the correction is subsequent to the photocopy. In this case then, the paratextual elements of TEI, in a rather paradoxical way, serve to come to assistance of that temporal dimension which the TEI-XML model implicitly rejects.

2.3.1 Encoding of A

```
<div type="second corrected version" n="M004b1_M004b2">
  <ab n="A">
    <seg type="l">Essere matita è segreta ambizione.<lb/></seg>
    <seg type="l">Bruciare sulla pagina lentamente<lb/></seg>
    <seg type="l">e nella pagina restare<lb/></seg>
    <seg type="l">in altra nuova forma suscitato.<lb/></seg>
    <seg type="l">Diventare così da carne segno, e da strumento<lb/></seg>
    <seg type="l">esile ossatura del pensiero,<lb/></seg>
    <seg type="l"><note type="annotation" resp="editor" xml:id="v2" next="v6">the correction to this part of the text is
      subsequent to the drafting of the sixth version of the poem</note><emph rend="square brackets"><del hand="M"
      type="overstrike">la sua rete di
    vene sottili</del></emph><lb/></seg>
  </ab>
</div>
```

2.3.2 Encoding of B

```
<div type="sixth version" n="M006_M007">
  <ab n="A" id="v6" prev="v2">
    <seg type="l">Essere matita è segreta ambizione.<lb/></seg>
    <seg type="l">Bruciare sulla pagina lentamente<lb/></seg>
```

```
<seg type="l">e nella pagina restare<lb/></seg>
<seg type="l">in altra nuova forma suscitato.<lb/></seg>
<seg type="l">Diventare così da carne segno, e da strumento<lb/></seg>
<seg type="l">esile ossatura del pensiero<lb/></seg>
<seg type="l">la sua rete di vene sottili.<lb/></seg>
</ab>
</div>
```

Conclusions: The Epistemology of Variation

In this paper we have described and analyzed some methods for representing textual phenomena of the writing process by means of the TEI-XML Guidelines. The term *epistemology of variation*,[14] invoked here for the first time, may help to express what is at stake in this representation: not simply an account of evidence whose significance begins and ends with the data, but an account of a system of knowledge in which the relation between states is as important as any state taken alone. As already mentioned, markup languages appear to be — at the moment — the most flexible scientific and economical solution for the representation of digital text. However, as already explained, it would be naive and probably counterproductive to claim that we are satisfied. Many (but not all) of the phenomena described here can be represented only with great effort or with an unacceptable level of imprecision. XML, like its predecessor SGML, was originally designed as an instrument for archiving and information retrieval, whereas the *mouvance* of the text poses challenges for the encoding and representation that can only be resolved by the development of a different model of encoding, and by the design of an adequate user interface on the application level. Nevertheless, it should be pointed out that without an adequate theoretical reflection on text in the digital era no instrument can offer satisfying solutions. The variant is "complex stratified knowledge" [Benozzo 2004, 52] which embodies an epistemological and cultural status. In fact, variation, from evolutionary biology [Edelman 2006] to cognitive anthropology [Sperber 1996], is at the heart of the phenomena and processes of diffusion of culture. In other words, a certain degree of instability is immanent to the transmission of knowledge. Variation therefore calls into question the notion of "repeatability," an important concept in disciplines concerned with the transmission of information — including philology. The "failure to repeat" [Ferrer 2002, 52] is thus an annoying spanner in the conceptual works of any procedure based on recursive structures. This happens because the variant, understood as a symptom of the processual (rather than the editorial) dimension is an "equally valid alternative," which does not expect the attainment of "truth." The new epistemological law revealed to us by the dynamics of the writing process is part of the actual paradigm-shift which concerns the sciences of language, where it emerges as a dialectic between the concept of "system" and that of "process" (or in our case between text and writing):

35

> Of particular interest appears to be the recent deepening of the notion of "process," which emerges in a separate but convergent manner from the study of phonetics and phonology, and from morphology and syntax ... These results converge towards a vision that is much more complex and empirically founded on the problem of the linearity of linguistic phenomena, showing that at each level of analysis global planes come into play behind the apparent seriality of linguistic production. ... One significant consequence of the creation of these new methods is that they succeed in reducing interest in the notion of process to the second level to the benefit of the notion of process. [Sornicola 2005, 37]

Even though it is always risky to propose analogies, this epistemological turn can be compared to that "second phase" in the history of physics, which is the subject of the work of the chemist and epistemologist Ilya Prigogine. For Prigogine post-Newtonian physics, having discovered disequilibrium, ceased to describe phenomena in terms of stability and uniformity, as Boltzmann had done, construing a model — whose results echo those of the forest of "bifid trees" discovered by Bédier — "from the irreversible evolution of the population of particles towards a state of equilibrium," which has the effect of describing, exactly as Bédier noted on the subject of the tradition of the *Lai de l'ombre* (Bédier 1929-1985), "an evolution towards uniformity" [Prigogine & Stengers 1988, 25–26]. At the centre of Prigogine's argument is the accusation directed against Newtonian physics that it ignores time: "...opposed to dynamic eternity is the second principle of thermodynamics... Physics was finally able to describe, like the other sciences, a world open to history." The second phase of physics "is that of the instability of the elementary particles and their complexity"; subsequently he adds "the discovery of structures not in equilibrium, which overturns the dogma that assimilates the

36

growth of entropy to molecular disorder" [Prigogine & Stengers 1988, 44]. This "non-equilibrium" greatly resembles the notion of process developed and analysed in various areas of the humanities: "If one can say, *a posteriori* , that this second period has seen the discovery of a world of processes of creation, far removed from the correct world of eternal laws, which characterised the ideal of classical physics" [Prigogine & Stengers 1988, 44]. The analogy could continue, for example, by noting that the philological concept of "original text" presupposes a dismissal of time, establishing a principle of reversability: apart from certain effects (the tradition) it is always possible to jump back to the cause (the original). But in both cases — in the physics of Boltzman and the philology of Lachmann — to go back to the cause is a metaphysical undertaking: texts and elementary particles are both ascribed to a world which cannot be interpreted or represented in genealogical terms or by following rigid principles of causality.

Applied to our case, that is, the representation and use of digital text, this vision of time, open to history, forces us to

37

1. acknowledge that the actual computational models, and in particular formal languages like XML, are not always adequate for the representation of unstable cultural information structures such as textual variants;
2. rethink the models of digital documents, incorporating the vision of text as a process, a plastic phenomenon, part of a temporal continuum in which all of its "phases and structures" (material or abstract[15]) have "intrinsic meaning." [16]

Addendum

The Encoding Model for Genetic Editions being proposed as a new chapter of the TEI Guidelines [Burnard et al. 2010] was published after this paper was accepted. Having read this document carefully, we still think that it does not address the essential difficulties we have raised in our study.

38

Acknowledgments

Notes

[1] Of the number of studies on the dynamics of the textual process (mostly modern and contemporary) the first proposals for computational solutions appeared at the beginning of the 90s; see in particular [Brockbank 1991], [Lebrave 1991], [Mordenti 1992] and [Ferrer 1995]. However, it is Jerome McGann [McGann 2001], perhaps more than any other scholar in the last few years, who has advanced the idea of text dynamics by developing a theoretical direction begun by Donald F. McKenzie [McKenzie 1986], and by applying it to important digital projects. Useful overviews of the connections between genetic criticism, Anglo-American textual bibliography and other European philological schools of thought are provided by [Morrás 1999], [Lernout 1996] and more recently by contributions to the first volumes of *Variants,* the review of the European Society for Textual Scholarship. For a reconstruction of the historic and theoretical links between Italian variant criticism and French genetic criticism see [Giaveri 1993], and the recent contributions in the international journal *Ecdotica*.

[2] On this basis, referring to Hans Zeller [Zeller 1975], Peter Shillingsburg and Jerome McGann (although omitting to mention Italian variant criticism and French genetic criticism), Susan Schreibman has coined the term *versioning* to describe the new editorial theory, which ought to be reserved for the development of digital *variorum* editions [Schreibman 2002, 288]. For a recent attempt to import French textual scholarship into the Anglo-American context see [Bushell 2007].

[3] Shillingsburg developed the concept of "script acts" mentioned above.

[4] French genetic criticism refers to a single or multiple authorial writing artifact as *avant-texte* (pre-text), while the entire corpus of manuscripts,

drafts, notes, etc., is termed *dossier genetique.*

[5] Even non-XML based markup systems like LMNL and TexMECS still model the text as a hierarchical structure (although not as a tree). In any case they are described as "purely experimental" [Johnsen and Huitfeldt 2007] and "under development" [Tennison et al. 2009] by their own authors, and so are not yet practical alternatives. COCOA, though often cited as allowing non-hierarchical structures, in fact imposes no real structure on the text other than to define a reference system [Hockey and Martin 1988].

[6] For a description of these materials see on the web site the introduction by Tommaso Lisa (now in [Lisa 2004, 9–23]).

[7] A TEI-XML encoding of all the writing drafts (autographs, typewritten and printed versions) of a corpus of approximately ten compositions by Valerio Magrelli is available in [Martiradonna 2004].

[8] For an analysis of the changes between versions P4 and P5 of the TEI Guidelines see Lou Burnard, "La TEI Lite dalla P4 alla P5: continuità e cambiamento," in [Burnard & Sperberg-McQueen 2005, 201–212].

[9] As expected by all markup languages derived from SGML (such as HTML), elements are required to start with an opening tag and end with a closing tag (the forward slash / indicates the closing of a tag). Since there is no room here for a precise description of TEI-XML we are constrained to refer to the manual of Burnard and Sperberg-McQueen [Burnard & Sperberg-McQueen 2005] or to the TEI website.

[10] For a discussion and a proposal for other hacks to avoid overlapping hierarchies, see for example [Bauman 2005a].

[11] Chap. 11 Representation of primary sources; Chap. 12 Critical Apparatus.

[12] It can be seen here how the act of encoding, by forcing us to represent through discrete elements the writing process continuum, could produce redundant structures. In fact, the registration of the writing stages includes also the already represented elements: see for example the "clone" of `<list>` (bottom list of adjectives), which we can find both in the second (`varSeq="2"`, r. 42) and third passage (`varSeq="3"`, r. 55) of the encoding of the fourth verse line (Table 1, lines 39-70). This redundancy reflects "spatially" the idea that each textual movement (each modification) should be considered as an independent system which, in its turn, needs a separate encoding/representation.

[13] Also where the empty element is necessary to avoid overlap.

[14] Although never explicitly cited, we recognize both by use of this term, and also through our first reflections on the concept, our debt to [Cerquiglini 1989].

[15] The phases of the process of writing are abstract and can be reproduced via digital "simulations"; but the structures of a document (the product) can be either material or abstract: the trace left on the page is a material fact, but the insertion of a title, the start of a new paragraph, the insertion of a note etc. exist as categories and as models which from time to time take concrete form in situations and expressions. As showed in researches that span from authorial medieval manuscripts [Barolini & Storey 2007] up to contemporary editions [Giuliani 2006], the text is the result of these material and abstract interacting forces, which are always mutually influencing one another.

[16] We are adapting to the domain of textual fluidity a formulation suggested by the second generation of cognitive linguists and expressed primarily in opposition to dualistic and formal approaches to language. According to cognitive linguist Langacker [Langacker 1987-1991] language cannot be considered a module separated from other mental activities and operations (cf. [Formigari 2001, 271]). In applying this conceptual model to the text, we could say that the abstract, material and procedural aspects of writing are all responsible for building the "meaning" of the text.

Works Cited

Austin 1976 Austin, J.L. *How to Do Things with Words: Second Edition*. Oxford: Oxford University Press, 1976.

Barolini & Storey 2007 Barolini, T., and H.W. Storey, eds. *Petrarch and the textual origins of interpretation*. Leiden; Boston: Brill, 2007.

Bart 2006 Bart, P.R. "Experimental markup in a TEI-conformant setting". *Digital Medievalist* 2: 1 (2006).

Bauman 2005a Bauman, Syd. "TEI HORSEing Around". Presented at *Extreme Markup Languages Conference* (2005). *Proceedings of the Extreme Markup Languages Conference* (2005). http://conferences.idealliance.org/extreme/html/2005/Bauman01/EML2005Bauman01.html.

Benozzo 2004 Benozzo, F. "Filologia al bivio: ecdotica celtica e romanza a confronto". *Ecdotica* 1 (2004), pp. 24-54.

Brockbank 1991 Brockbank, P. "Towards a Mobile Text". In I. Small and M. Walsh, eds., *The theory and practice of text-editing. Essays in honour of James T. Boulton*. Cambridge: Cambridge University Press, 1991. pp. 90-106.

Burnard & Bauman 2007 Burnard, Lou, and Syd Bauman, eds. *TEI P5: Guidelines for electronic text encoding and interchange*. Text Encoding Initiative Consortium, 2007. http://www.tei-c.org/Guidelines/P5/.

Burnard & Sperberg-McQueen 2005 Burnard, L., and M. Sperberg-McQueen. *Il manuale TEI Lite. Introduzione alla codifica elettronica dei testi letterari (a cura di F. Ciotti)*. Milano: Edizioni Sylvestre Bonnard, 2005.

Burnard et al. 2010 Burnard, L., F. Jannidis, E. Pierazzo and M. Rehbein. *An Encoding Model for Genetic Editions*. http://users.ox.ac.uk/~lou/wip/geneticTEI.doc.html.

Bushell 2007 Bushell, S. "Textual process and the denial of origins". *Textual Cultures* 2: 2 (2007), pp. 100-117.

Buzzetti 2002 Buzzetti, D. "Digital Representation and the Text Model". *New Literary History* 33 (2002), pp. 61-88.

Cerquiglini 1989 Cerquiglini, B. *Éloge de la variante. Histoire critique de la philologie*. Paris: Seuil, 1989.

Ciotti 2005 Ciotti, F. "La codifica del testo, XML e la Text Encoding Initiative". In L. Burnard and M. Sperberg-McQueen, eds., *Il manuale TEI Lite: introduzione alla codifica elettronica dei testi letterari*. Milano: Sylvestre Bonnard: , 2005. pp. 9-42.

Edelman 2006 Edelman, G.M. *Second nature: brain science and human knowledge*. New Haven: Yale University Press, 2006.

Ferrer 1995 Ferrer, D. "Hypertextual Representation of Literary Working Papers". *Literary and Linguistic Computing* 10: 2 (1995), pp. 143-145.

Ferrer 2002 Ferrer, D. "Production, Invention and Reproduction: Genetic vs. Textual Criticism". In E. Bergmann Loizeaux and N. Fraistat, eds., *Reimagining Textuality: Textual Studies in the Late Age of Print*. Madison: University of Wisconsin Press, 2002. pp. 48-59.

Formigari 2001 Formigari, L. *Il linguaggio. Storia delle idee*. Roma-Bari: Laterza, 2001.

Giaveri 1993 Giaveri, M.T. "La critique génétique en Italie. Contini, Croce et l'étude des paperasses". *Genesis* 3 (1993), pp. 9-29.

Gigliozzi 1998 Gigliozzi, G. "La galassia Von Neumann: il testo fra piombo e byte". Presented at *I nuovi orizzonti della filologia. Ecdotica, critica testuale, editoria scientifica e mezzi informatici elettronici*, sponsored by Atti del convegno Internazionale dell'Accademia Nazionale dei Lincei in collaborazione con l'Associazione Internazionale per gli Studi di Lingua e Letteratura Italiana, Roma. Roma: Accademia Nazionale dei Lincei (27-29 maggio 1998).

Giuliani 2006 Giuliani, L., H. Brinkman, G. Lernout and M. Mathijsen, eds. *Texts in multiple versions: histories of editions, Variants*. Amsterdam-New York: Rodopi, 2006.

Hockey and Martin 1988 Hockey, S., and J. Martin. *The Oxford Concordance Program*. Oxford: Oxford University Computing Service.

Johnsen and Huitfeldt 2007 Johnsen, and Huitfeldt. "Preserving Information About Linearization in Document Graphs". In S. Schmidt S. Siemens, R. Siemens, A. Kumar and J. Unsworth, eds., *Digital Humanities 2007 Conference Abstracts*. 2007. pp. 104-106.

Langacker 1987-1991 Langacker, R.W. *Foundations of cognitive grammar*. Stanford: Stanford University Press, 1991.

Lebrave 1991 Lebrave, J.-L. "L'hypertexte et l'avant-texte". In J. Anish and J.-L. Lebrave, eds., *Text et Ordinateur. Les Mutations du Lire-Écrire. Actes du colloque interdisciplinaire tenu à l'Université de Paris X, Nanterre, 6-7-8 juin 1990*. Paris: Éditions de l'Espace Européen. pp. 101-117.

Lernout 1996 Lernout, G. "Anglo-American Textual Criticism and the Case of Hans Walter Gabler's Edition of Ulysses". *Genesis* 9 (1996), pp. 45-65.

Lisa 2004 Lisa, T. *Scritture del riconoscimento. Su Ora serrata retinae di Valerio Magrelli*. Roma: Bulzoni, 2004.

Magrelli 1996 Magrelli, V. *Poesie (1980-1992) e altre poesie*. Torino: Einaudi, 1996.

Martiradonna 2004 Martiradonna, V. *La codifica elettonica dei testi. Un caso di studio*. Thesis, Tesi di laurea in Lettere, Facoltà di Scienze Umanistiche, Università di Roma La Sapienza: 2003-2004.

McGann 2001 McGann, Jerome. *Radiant Textuality: Literature After the World Wide Web*. New York: Palgrave Macmillan, 2001.

McKenzie 1986 McKenzie, D.F. *Bibliography and the Sociology of Texts. The Panizzi Lectures, 1985*. London: The British Library, 1986.

Mordenti 1992 Mordenti, R. "Informatica e filologia". In *Calcolatori e Scienze Umane. Scritti del convegno organizzato dall'Accademia dei Lincei e dalla Fondazione IBM Italia*. Milano: Fondazione IBM Italia e Etas Libri, 1992. pp. 236-272.

Morrás 1999 Morrás, M. "Informática y crítica textual: realidades y deseos". In J.M. Blecua G. Claveria, C. Sánchez and J. Torruella, eds., *Filología e informática. Nuevas tecnologías en los estudios filológicos*. Barcelona: Editorial Milenio-Universidad Autónoma de Barcelona, 1999. pp. 189-210.

Pierazzo 2007 Pierazzo, Elena. "The Encoding of Time in Manuscripts Transcription: Toward Genetic Digital Editions". Presented at *Digital Humanities 2007. Digital Humanities 2007 Conference Abstracts* (2007).

Prigogine & Stengers 1988 Prigogine, I., and I. Stengers. *Entre le temps et l'éternité*. Paris: Librairie Arthème Fayard, 1988.

Renear 2005 Renear, A. "Text from Several Different Perspectives: The Role of Context in Markup Semantics". In C. Nicolás Martínez and M. Moneglia, eds., *Computers, Literature and Philology. CLIP 2003. La gestione unitaria dell'eredità culturale multilingue europea e la sua diffusione in rete*. Firenze: Firenze University Press, 2003. pp. 25-34.

Schreibman 2002 Schreibman, Susan. "Computer-mediated Texts and Textuality: Theory and Practice". *Computers and the Humanities* 36: 1 (2002), pp. 283-293.

Shillingsburg 2006a Shillingsburg, P. "Verso una teoria degli atti di scrittura". *Ecdotica* 2 (2006), pp. 60-79.

Sornicola 2005 Sornicola, R. "Italiano parlato, dialetto parlato, parlato". In E. Burr, ed., *Tradizione e Innovazione. Linguistica e filologia italiana alle soglie dì un nuovo millennio, Atti del VI Convegno SILFI (Duisburg, 23 Giugno - 2 Luglio 2000)*. Firenze: Franco Cesati, 2000. pp. 25-39.

Sperber 1996 Sperber, D. *Explaining culture: a naturalistic approach*. Oxford: Blackwell, 1996.

Tennison et al. 2009 Tennison, J., W. Piez, J. Cowan and M. Palmer. *The LMNL Wiki*. http://www.lmnl.org/wiki/index.php/Main_Page.

Vanhoutte 2003 Vanhoutte, E. "Display or Argument: Markup Visualisation for Electronic Scholarly Editions". In T. Burch J. Fournier, K. Gärtner and A. Rapp, eds., *Standards und Methoden der Volltextdigitalisierung. Beiträge des Internationalen Kolloquiums an der Universität Trier, 8.-9. Oktober 2001*. Stuttgart: Franz Steiner Verlag, 2003. pp. 71-96.

Vanhoutte 2006 Vanhoutte, E. "Prose Fiction and Modern Manuscripts: Limitations and Possibilities of Text-Encoding for Electronic Editions". In L. Burnard K. O'Brien O'Keefe and J. Unsworth, eds., *Electronic Textual Editing*. New York: Modern Language Association of America, 2006. pp. 161-180.

Vetter and McDonald 2003 Vetter, L., and J. McDonald. "Witnessing Dickinson's Witnesses". *Literary and Linguistic Computing* 18: 2 (2003), pp. 151-165.

Zeller 1975 Zeller, H. "A New Approach to the Critical Constitution of Literary Texts". *Studies in Bibliography* 28 (1975), pp. 231-264.