

## Prose Fiction and Modern Manuscripts

### *Limitations and Possibilities of Text-Encoding for Electronic Editions*

**Edward Vanhoutte**

*Centrum voor Teksteditie en Bronnenstudie - Centre for Scholarly Editing and Document Studies (Royal Academy of Dutch Language and Literature), Belgium*

*Thinking things out involves saying things to oneself,  
or to one's other companions, with an instructive intent.<sup>1</sup>*

(Gilbert Ryle, *The Concept of Mind*, 294-295)

The hypertext edition, the hypermedia edition, the multimedia edition, the computer edition, the digital edition, the electronic edition are all synonymous labels for a concept without a definition. The use of “edition” in these labels presuppose a conventional understanding of that word, and is based on the implicit assumption that there is a general agreement on what an edition is. But there is no such thing. In textual studies we certainly mean “scholarly” edition whereas in ‘strict’ bibliography “the term ‘edition’ refers quite properly to the identifiably separate type-setting of *any* book, whether it is an edition in the text-critical sense or not” (Greetham 348). Any kind of available text, therefore, qualifies as an “edition”, and any kind of electronically available text qualifies to be labelled an “electronic edition”, just as any printed text can be called a “paper” or a “print edition”. The words “digital” and “electronic” only describe the materiality (as opposed to “print”) by which data comes before the reader, “computer” refers to the distribution and access medium (as opposed to “book” or “codex”), and “hypertext”, “hypermedia”, and “multimedia” hint at the functionality of the electronic product (as opposed to “linear”). As John Lavagnino pointed out in a report on a lecture I gave in London in 2000,<sup>2</sup> “Ten years ago, it seemed sufficient to say that you were going to create such an edition in the form of a hypertext: often with very little elaboration on just what the result would be or why it would be significant, as though the medium itself would automatically make such an edition significant.”

To allow a functional debate on editing and editions in the electronic paradigm, editors should provide an explicit definition of an electronic edition as well as the kind of scholarly edition they are presenting in electronic form. They should do so not only by using qualifying adjectives or labels in the titles or subtitles of their electronic products, for which conventional terminology could be used where applicable, but also by means of a consistent argument in a textual essay that accompanies the edition proper.

This is one of the six requirements I formulate in my working definition of an electronic scholarly edition which is proposed and discussed in the first section of this essay. As the attentive reader will notice, the formulation and discussion of this definition are influenced by an interest in and concern for non-critical editing and genetic transcription of modern manuscript material, which will be two of the main topics covered in this essay. The case study of my own electronic-critical edition of Stijn Streuvels' *De teleurgang van den Waterhoek* in the second section will serve as a test case for my definition, and an introduction to some limitations and possibilities of text encoding for electronic editions of modern prose texts. The assessment of this edition will be used as a starting point for elaborations on issues concerned with the production of literal (non-critical) transcriptions and genetic editions of modern manuscript material with the use of text-encoding as proposed by the Text Encoding Initiative (TEI). In the course of the third section, all these issues will be covered and the French school of *critique génétique* that has developed a decade-long theory and practice of dealing with modern manuscript texts in a specific way will be introduced. The attempt at applying their central interest in the internal dynamics of the modern manuscript and the study of the genetic continuum in a so-called *dossier génétique* to the transcription of manuscript material by means of text-encoding, is the focus of this third section of this essay: the current inability to encode overlapping hierarchies and time (both absolute and relational) elegantly in (non-critical) manuscript transcriptions will take up most of this section. I finish with some concluding remarks.

## 1. Definition and aims of an electronic edition

My full working definition of an electronic (scholarly) edition has six parts. By electronic edition, I mean 1. the immediate result or some kind of spin-off product from textual scholarship, 2. which is intended for a specific audience and designed according to project-specific purposes, 3. which represents at least one version of the text or the work, 4. which has been processed from a platform-independent and non-proprietary basis that is both stored for archival purposes and is made available for further research (Open Source Policy), 5. whose creation is documented as part of the edition, and 6. whose editorial status is explicitly articulated in the edition.

With respect to this definition, five immediate observations have to be made. First, by defining an electronic edition as a “result” or a “product” which has been “processed”, I plead for a practice in which the construction of a digital archive which contains all data (encoded transcriptions, high resolution image files etc.) differs from and precedes the generation of the edition. I have called this the *Archive/Museum model*<sup>3</sup> (Vanhoutte, *Resistance*), Espen Ore speaks in this case of *Teksttilretteleggelse/arkivoppbygging* and *Utgaveproduksjon* (Ore, *Publisering* 143). Second, in so far that I define an electronic edition in productive terms, it is indeed true, as G. Thomas Tanselle writes in “Textual Criticism at the Millenium” paraphrasing Peter Shillingsburg, that “an electronic edition is a form of presentation and, as such, does not pose a different set of theoretical issues from the one faced by editors who present their work in a different form.”(33) What Tanselle and Shillingsburg here seem to overlook is that the practice of creating an edition with the use of text-encoding calls for explicit ontologies and theories of the text which do generate new sets of theoretical issues.<sup>4</sup> Maybe they are not different sets of editorial issues, but they are certainly new sets of textual issues such as problems of document architecture, encoding time etc. Third, I do not consider it a requirement for an electronic edition to display textual variation in an apparatus or in any other way, for in my view this is a project-specific purpose. However, it follows from the first requirement that the study of textual variation—where it appears—is an essential part of the research involved in creating an electronic edition. Even a reading edition cannot be made without a serious examination of the textual variants. Fourth, the requirement for the

presentation of at least one version of the text or the work is no fundamental call for a critically established editor's text. Diplomatic transcript editions (Greetham 391) or facsimile editions, for instance, do not present a critically edited text, but they should not pretend to be non-editorial. The editor is always present in the organisation of the material and the transcription of source documents. Point 3 in my definition specifies that a database of textual variation cannot be considered an electronic edition. If one is forced to choose between an edited text or textual variation, the edited text must always take priority. The category of electronic edition includes more than critical, diplomatic, facsimile, and reading editions. Mixed flavours editions (which have characteristics of all of them but are none of them exclusively), as well as archives which present editions, and editions which are organized as archives all qualify as electronic editions.<sup>5</sup> Fifth, the working definition does not prescribe any hypertextual features to be included in the edition, because the emphasis on hypertext in the current debate on electronic editing is still often beside the point and is about to outgrow its hype. If the use of hypertext does not add any fundamental advantages to the electronic edition over the codex based edition, it's better to stick to the book. Indeed, hypertext is just the visualization of linking which DeRose & Van Dam define as "the ability to express relationships between places in a universe of information"<sup>6</sup> (9), and which should explicitly be marked or should be generated automatically by making use of some sort of markup and path scheme. Consequently, the syntax of this markup and the markup-language becomes essential in the design of an electronic edition with hypertext functionality. The presence of hypertext functionality in an electronic text, however, does not guarantee its scholarly integrity.

As G. Thomas Tanselle defined in his essay "The Varieties of Scholarly Editing", scholarly editing is the considered act of reproducing or altering texts. Whether the result of that act is published in print or as an electronic product does not matter to this general definition. Nor does it affect its primary aims. Above all, the scholarly edition is aimed at articulating the editors' notions, perspectives, or theories of the texts—or of what the texts should have been in the case of critical editions. By doing so, the scholarly edition could invoke the literary debate by providing the materials and tools to explore new ways of understanding and studying the text. Only the format of these materials and tools will differ in the print and

electronic edition, not their intentions. The ultimate, and for technical reasons the most problematic, aim however is to preserve our cultural heritage.

## 2. A case study

My electronic edition of the classic Flemish novel *De teleurgang van den Waterhoek*<sup>7</sup> by the Flemish author Stijn Streuvels<sup>8</sup> (1871-1969)—the first electronic scholarly edition of any sort of a Dutch/Flemish literary work—was published as an “electronic-critical edition” on CD-ROM by Amsterdam University Press in 2000. Unfortunately, at the time of publication, however, I had not defined ‘electronic edition’ in the general introduction, neither had I spelled out the ways in which the edition was “critical” nor provided an account of the *linkemic* approach towards textual variation and the eclecticism towards conventional editorial theory the editors had taken.<sup>9</sup> As a result, in the absence of these explicit statements, some users and reviewers misunderstood the edition. Some criticized as disorienting, for instance, the presentation of two critical texts, the mixed use of text-transcriptions and digital facsimiles as representations of the documentary source material, or the absence of an apparatus variorum, even though these features were actually the three pivotal theoretical statements of the edition. The following short survey of the edition’s main editorial procedures and problems and discussion of both the solutions suggested by the edition and the questions it raises will serve, I hope, as an object lesson and case study,<sup>10</sup> by introducing the issue of non-critical and genetic transcription/editing of modern manuscript material.

### 2.1. Editorial Principles and Markup

In *Scholarly Editing in the Computer Age*, Peter Shillingsburg defines five formal orientations of editing: documentary, aesthetic, authorial, sociological, and bibliographic (15-27). Except for the aesthetic orientation, all of them were applied to editing *De teleurgang van den Waterhoek*.<sup>11</sup> The aim of the editorial project was to explore different ways to deal with textual instability, textual variation, the genetic reconstruction of the writing process, and the constitution of a critically restored reading text. Whereas the reading edition in bookform which was published together with the electronic edition presents one text (the edited text of

the first print edition),<sup>12</sup> this seemingly best-text approach is countered by the electronic-critical edition, where instability and *versioning*<sup>13</sup> are the governing principles in the presentation of many texts.

On the basis of their documentary status and appearance, the many texts in the edition could be divided into two groups:

*Complex Documentary Sources*

- the complete holograph manuscript (1927)

- the corrected authorial copy of the prepublication which served as a printer's copy for typesetting the first print edition (1927)

- the corrected author's copy of the first print edition which was used as printer's copy for typesetting the second print edition (1939)

*Simple Documentary Sources*

- the prepublication of the novel which was published in installments in the Dutch literary journal *De Gids* (1927)

- the first print edition (1927)

- the revised second print edition (1939)

In the electronic edition, the documentary sources of the complex group are represented by their digital facsimiles, whereas the texts of the simple group are TEI compliant full text representations. By including either a full text version or a digital facsimile version of a documentary source, the electronic edition shows a mixed approach towards Peter Shillingsburg's (and Thomas Tanselle's) first requirement to provide both a "a full accurate transcription" and a "full digital image of each source edition" in an electronic edition. (Shillingsburg, *Principles* 28). This mixed approach, however, is partly compensated for by the possibility of deducing the physical form of the prepublication from the facsimiles of the

corrected prepublication, and the first print edition from the corrected first print edition. Only the presentations of the holograph manuscript and the second print editions do not have a full text and a facsimile counterpart respectively.

For sociological reasons, the edition presents two critically edited texts: the versions of the first and the second print editions. The first print edition constitutes an important moment in the genetic history of the novel. It is the first finished version of the text—for we consider the prepublication as one step in the writing process leading towards the first print edition—and its reception by literary critics is the reason for the drastic revision of the text in preparation of the second print edition. This second print edition is sociologically equally important because it is the version that generations of readers have read and which the author constituted in interaction with contemporary society.<sup>14</sup> In constituting these texts, we have applied the principles of the German (authorial) editorial tradition which only allows justified corrections of manifest mistakes. In these two critical texts, the emendations<sup>15</sup> were documented by the use of the <CORR> element, containing the correction, and a SIC attribute whose value documents the original reading. The editor responsible for the correction (in this case EV for Edward Vanhoutte) was specified in a RESP attribute: <CORR SIC="katin" RESP="EV">kattin</CORR>. The text of the prepublication was retained for historical reasons, and the inverse markup was used. The uncertain readings<sup>16</sup> were encoded with the aid of the <SIC> element, containing a suggested correction in a CORR attribute: <SIC CORR="kattin">katin</SIC>.<sup>17</sup> Although both systems are equivalent for the computer, they do articulate different views on the text. In the corresponding use of the inverse markup, one can see the thin line between critical editing and non-critical or documentary editing.

## ***2.2. A linkemic approach to textual variation***

The electronic edition of *De teleurgang van den Waterhoek* contains four sections:<sup>18</sup> 1. an account of the editorial principles, consisting of a genetic history of the work, an account of the constitution of the texts, and a description of the transmission and publication history, 2. the actual edition which presents the textual variation against the orientation text by making

use of the *linkemic* approach, 3. the separate documentary sources, and 4. the diplomatic edition of 71 relevant letters for the reconstruction of the textual history, selected from the correspondence of the author.<sup>19</sup>

The third section of the edition, documentary sources, presents the six versions of the text chronologically, allowing the user to consult each of them separately. It is of course true that “every form of reproduction can lie, by providing a range of possibilities for interpretation that is different from the one offered by the original” (Tanselle, *Reproductions* 33) and the process of imaging is a process of interpretation.<sup>20</sup> In order for the user of the edition to be able to evaluate what she sees, the facsimile editions are accompanied by a full account of the imaging procedure including the documentation on the soft- and hardware (and settings) used in the project, which I believe is an essential requirement. No facsimile can of course substitute the original, but it is the best approximation we can offer the interested user.

The user of the edition can also read all six versions together in the second section. This part of the electronic-critical edition presents the edited text of the first print edition as the *orientation text* around which the hypertext presentation of textual variation is organized. Instead of linking the orientation text to an apparatus variorum, the editors opted for what I have called a *linkemic* approach to textual variation. I define a *linkeme* as the smallest unit of linking in a given paradigm. This unit can be structural (word, verse, sentence, stanza, etc.) or semantic. In the case of the glossary provided with the orientation text, the *linkeme* is of a semantic class which can be defined as “the unit of language that needs explanation”. In the case of the presentation of textual variation, the *linkeme* is a structural unit, namely the paragraph. In the actual hypertext edition it is possible to display all the variants of each paragraph in all six of the versions on the screen. This is made possible by a complicated architecture on the code side which allows for hypertext visualisation on the browser side. This architecture was *automagically* generated from the digital archive by a suite of SED and AWK scripts. This linkemic approach provides the user with enough contextual information to study the genetic history of the text, and introduces new ways of reading the edition. Because of the fact that a new document window, displaying a text version of the user’s choice, can be opened alongside the hypertext edition, every user can decide on which text to read as her own base text. The hypertext edition can then be used as a sort of apparatus with

any of the versions included in the edition. This way, hypertext and the linkemic approach enable the reading and study of multiple texts and corroborate the case for textual qualifications such as variation, instability and genetic (ontologic/teleologic) dynamism.

### 3. Modern manuscripts

As we will see, despite its strengths this practice is problematic for a genetic edition based on modern manuscript material. In what follows, I will start with two observations which are inherently subjective and personal, but nonetheless—I think—true. The reality they're signalling I will then confront with the reality of transcribing modern manuscripts out of genetic interest, which will lead to a brief treatment of the goals of the mainly French school of *Critique Génétique*, the peculiarities of the modern manuscript and the problems resulting from trying to achieve the work with the current TEI proposals. As an alternative I will suggest that further research on a methodology and practice of non-critical editing or transcription of modern manuscript material may result in markup strategies which could be applied to the constitution, reading and analysis of a so-called *dossier génétique*. My approach to the manuscript as a filtered materialization of an internal creative process which can be compared to the process of internal monologue or dialogue and can thus be considered as a form of speech, might be helpful in this respect.

#### 3.1. Observation 1: (Non-)Critical editing

In his article “Literal transcription—Can the text ontologist help” Allen Renear points out that the distribution of attention to and resources for the study of *critical* editing and the study of *non-critical* editing is in inverse proportion to their relative practical importance (25). Apart from Mary-Jo Kline's very useful *A Guide to Documentary Editing*, it is for instance very difficult to find an extensive and coherent treatment of *non-critical* editing in handbooks and survey articles on scholarly editing, and the specialized journals mainly if not exclusively focus on the theory and practice of *critical* editing (24). The eight steps Wilhelm Ott outlined for the production of a critical edition in his “Computers and Textual Editing”, and which Susan Hockey took as the framework for her chapter on “Textual Criticism and Electronic Editions”

in *Electronic Texts in the Humanities* (124-145), for instance, jump from the collection of witnesses (step 1) to the collation of the witnesses (step 2), and further to

- S the evaluation of the results of the collation (step 3)
- S the constitution of a copy text (step 4)
- S the compilation of apparatuses (step 5)
- S the preparation of indexes (step 6)
- S the preparation of the printer's copy (step 7)
- S and finally the publication of the edition (step 8).

The essential, difficult and time consuming step of the transcription of the primary textual sources is not explicitly mentioned in this outline. Instead it is silently folded into the second step i.e. the collation of the witnesses, and thus primarily oriented towards the representations or captions of the linguistic text whose collations produce a record of the inter-documentary variation. Optimally, transcription should precede the second step, as transcription or *non-critical* editing is nonetheless the principal activity of critical editing and the epistemic foundation for all further textual criticism and textual editing. Only on the basis of such transcriptions can the editor proceed to automatic collation, stemma (re)construction, the creation of (cumulative) indexes and concordances etc. by the computer.

This neglect of non-critical editing in otherwise instructive outlines for the textual critic is the more remarkable because in the not so distant past, as we may recall from reading for instance Ben Ross Schneider's report on the *London Stage Information Bank* in his *Travels in Computerland*, up to half the time, the efforts, and the budget of an electronic project were devoted to converting textual data into another format, in this case a machine-readable format. Electronic non-critical editing is concerned with the twofold transformation from one format into another: First the transformation from the text of a physical document to the transcription of that text, and second the transformation from one medium, the manuscript, to another, the machine-readable transcription.

Peter Robinson makes the important point about transcribing the *Canterbury Tales* manuscripts that there is no division between transcription and editing. "To transcribe a manuscript is to select, to amalgamate, to divide, to ignore, to highlight, to edit." (Robinson,

*Politics* 10). And if we consider editing as short for *scholarly* editing, then both *critical* and *non-critical* editing should always have a scholarly result. The reason for the neglect of non-critical editing in the theory and practice of textual criticism, however, is frequently, as I will argue in a moment, the lack of a satisfactory ontology of the text on which a methodology of non-critical editing could be modelled.

In contrast with the production scheme for critical editions I have just discussed, the *TEI Guidelines* do seem to make the explicit distinction between *non-critical* and *critical* editing by devoting two different—but consecutive—chapters on “Transcription of Primary Sources” (chapter 18) and “Critical Apparatus” (chapter 19). Although the former chapter starts with explaining that “It is expected that this tag set will also be useful in the preparation of critical editions, but the tag set defined here is distinct from that defined in chapter 19 *Critical Apparatus*, and may be used independently of it.” (453), the DTD-subset for transcription also allows the scholar “to include other editorial material within transcriptions, such as comments on the status or possible origin of particular readings, corrections, or text supplied to fill lacunae” (453) which are strictly speaking no features of literal transcription. So, the chapter that seemingly deals with non-critical editing in the *TEI Guidelines* addresses issues which are central in critical editing, and includes in its DTD-subset tags to encode them. A provocative reading of the rhetoric of the opening paragraphs of chapter 18 could then be paraphrased as “the scholar who wants to transcribe primary source material in a non-critical way, in fact only uses a small portion of the DTD-subset and does not exploit its capabilities to the full”. This is a typical signifier of the influence of the traditional schools of textual editing on the work of the TEI. Whereas they overemphasize the allegedly unimportance of non-critical editing in their theories, the French school of *critique génétique* mainly works with non-critical representations of the documents under study.

### **3.2. Observation 2: Methodology**

According to Allen Renear “Non-Critical editions ‘transcribe’, ‘reproduce’, ‘present’—all these words are used—the text of a particular document, with no changes, no subtractions, no additions. It gives us the text, the whole text and nothing but the text of an actual physical

document.” (24). However, as I have already suggested, an answer to the basic question of what text is, and hence what to transcribe is a prerequisite to the transcription of a primary source. Only when a project has a clear agreement on the ontology of the text, can a methodology for text transcription be developed.

Although the TEI subset for the transcription of primary source material “have not proved entirely satisfactorily” for a number of problems (Driscoll 81), they do provide an extremely rich set of mechanisms for the encoding of older texts and documents with a fairly “neat” appearance such as print editions, which are fairly static and stable. Although the chapter on transcription does quote examples from modern manuscript material, the result of the successful use of the TEI in critical and non-critical editing can mainly be observed in projects which deal with older material which are interested in the record of inter-documentary variation.

The transcription of modern manuscript material using TEI proves to be of a more problematic nature because of at least two essential characteristics of such complex source material: namely the notions of *time* and of *overlapping hierarchies*. Since SGML (and thus XML) was devised on the assumptions that a document is a logical construct that contains one or more trees of elements that make up the documents content (Goldfarb 18) several scholars began to theorize about the assumption that text is an ordered hierarchy of content objects (OHCO thesis), which always nest properly and never overlap,<sup>21</sup> and the difficulties attached to this claim.<sup>22</sup> The *TEI Guidelines* propose five possible methods to handle non-nesting information,<sup>23</sup> but state that:

Non-nesting information poses fundamental problems for any encoding scheme, and it must be stated at the outset that no solution has yet been suggested which combines all the desirable attributes of formal simplicity, capacity to represent all occurring or imaginable kinds of structures, suitability for formal or mechanical validation, and clear identity with the notations needed for simpler cases (i.e. cases where the textual features do nest properly). The representation of non-hierarchical information is thus necessarily a matter of choices among alternatives, of tradeoffs between various sets of different advantages and disadvantages.<sup>24</sup>

The editor using an encoding scheme for the transmission of any feature of a modern manuscript text to a machine-readable format, is essentially confronted with the dynamic concept of *time* which constitutes non-hierarchical information. Whereas the simple representation of a prose text can be thought of as a logical tree of hierarchical and structural elements such as book, part, chapter, and paragraph, and an alternative tree of hierarchical and physical elements such as volume, page, column, and line—structures which can be applied to the wide majority of printed texts and classical and medieval manuscripts—the modern manuscript shows a much more complicated web of interwoven and overlapping relationships of elements and structures.

Modern manuscripts, as Almuth Grésillon defines them, are “manuscripts which are considered as forming part of the genesis of a text, evidence of which evidence is given by several successive witnesses, and which show the writing process of an author. (244)<sup>25</sup> Therefore, the structural unit of a modern manuscript is not the paragraph, nor the page or the chapter, but the temporal unit of writing. These units form a complex network which are often not bound to the chronology of the page.

The current inability to encode these temporal and genetic features of the manuscript and the overlapping hierarchies with a single, elegant encoding scheme, forces the editor to make choices which result in impoverished and partial representations of the complex documentary source. Further, if such an encoding scheme would exist, and genetic transcriptions could be produced, the current collation software would need to be redesigned in order to take these non-nesting structures into account. Therefore, in the electronic edition of *De teleurgang van den Waterhoek*, we have opted to represent the complex documentary sources by means of digital facsimile's only, preserving the genetic context of dynamic writing process of the author. By mounting these digital images in a hypertext structure which confronts them with representations of the other witnesses in the edition, the instability of the text is emphasized, and the user is provided with a tool to reconstruct the writing process of the novel under study. This focus on hypertext functionality and image based editing<sup>26</sup> could give the false impression that transcriptions become superfluous, and by astounding their audience with digital facsimiles, clickable image maps, and javascript driven dynamic views

of the digital images, the developers of such editions very cunningly avoid having to include a full transcription of the documentary sources—and I have been one of them.<sup>27</sup>

A further example of this ‘fear’ to test systems against modern manuscript material is illustrated in a recent article by Eric Lecolinet, Laurent Robert and François Role on a sophisticated tool for the *Text-image coupling for Editing Literary Sources*. The article, which appeared in the thematic issue on *Image-based Humanities Computing of Computers and the Humanities* (Kirschenbaum) presents, I quote from the abstract,

a system devoted to the editing and browsing of complex literary hypermedia including original manuscript documents and other handwritten sources. Editing capabilities allow the user to transcribe images in an interactive way and to encode the resulting textual representation by means of a logical markup language[...] (Lecolinet et al. 49)

The first figure in this article shows a complicated page in terms of authorial interventions from Flaubert’s manuscript, but the five further figures which demonstrate the system at work, show a simple medieval manuscript and a simple handwritten text neither of which contain any authorial additions, deletions, substitutions or revisions. It seems that the texts used for demonstration have been chosen in function of the limited capabilities of the demonstrated system.

The mentioned ‘fear’ to test existing transcription systems with modern manuscript material of a complicated nature in several projects so far may signal the fact that a coherent system or methodology for the transcription of modern material still has to be developed and tested, and that an ontology of the text must be agreed on.

### ***Intermezzo***

“There has never been a single standard convention for the transcription of manuscript texts, and it is not likely that there ever will be one, given the great variety of textual complications that manuscripts—from all times and places—can present.” (Vander Meulen & Tanselle 201)

### 3.3. *Genetic Criticism - Critique Génétique*<sup>28</sup>

At the end of his previously mentioned article, Allen Renear announces that there are major challenges to his view of text ontology which sees the literal transcription as a “literal representation of the linguistic text of a particular document.” (29) Therefore, he adds, “it presents the author’s linguistic achievements, not the author’s linguistic intentions.” One such challenge, he sees in Jerome McGann’s identification and treatment of ‘bibliographic codes’ (30). *Critique Génétique* is yet another major challenge to his view of text ontology, in that it is also interested in things such as time, ductus, and topology of the page which are strictly speaking no parts of the text.

The problematic position of *critique génétique* amongst the other schools of textual criticism and textual editing is that its primary aim is to study the *pre-text (avant-texte)*, not so much as the basis to set out editorial principles for textual representation, but as a means to understand the genesis of the literary work. Therefore, *critique génétique* does not aim to reconstitute the optimal text of a work, and is eventually not interested in the text, but in the dynamic writing process which can be reconstructed by close study of the extant drafts, carnets, etc. Moreover, as Tanselle points out in a discussion of Antoine Compagnon's introduction to a special issue of *Romanic Review*<sup>29</sup>:

French genetic critics are generally opposed to the construction of editions, on the grounds that an apparatus of variants, derived from the classical model in which variants are departures from an author's final text, is inappropriate for an authorial *avant-texte* and implies a subordination of it. (*Millenium*, 27)

Rather than producing editions, the *généticiens* constitute a *dossier génétique* by localising and dating, ordering, deciphering, and transcribing all pre-text witnesses. Only then can they read and interpret the *dossier génétique*. This does not mean, however, that the publication of genetic editions is not possible. The influential French *généticien* Pierre-Marc de Biassi defines three different categories of genetic editions (Van Hulle 375-376):

1. *Transversal edition*: attempts to render “works” that were left unfinished because of the author’s sudden death or for whatever other reason.

2. *Horizontal edition*: reconstructs one particular phase in the writing process, e.g. the author's notebooks of a certain period.
3. *Vertical edition*: reconstitutes the complete textual history

The computer has been used both in trying to publish a genetic edition, and in constituting and reading a *dossier génétique*. The concept of hypertext has been used extensively to regroup a series of documents which are akin to each other on the basis of resemblance or difference in multiple ways but the several endeavours to produce such hypertext editions or dossiers, are too much oriented towards display. In the past as well as nowadays commercial software packages such as Hypercard, Toolbook, Macromedia, Adobe Acrobat and even Powerpoint have been used for genetic purposes, resulting in proprietary formats. Where standards for markup have been used, we see HTML as the unbeaten champion in products in which every link, colour, frame and animation has been hardcoded without a repository of SGML/XML based transcriptions of the manuscript texts. This is partly, as we have seen, because of a lacking or specifically non-text based ontology of the text.

### ***Intermezzo***

“The institutional and technical machinery of textual genetics should not blind us to the fact that the object that it purports to study will almost by definition escape ‘science’.

What textual genetics studies is in effect something that cannot be observed, that cannot become an object: the origin itself of the literary work.” (Laurent Jenny)<sup>30</sup>

### ***3.4. Putting time back in manuscripts***

I want to propose a methodology which might help us in combining the study of “what cannot be observed” in very observable markup.

I am arguing that a writing process, and hence any text resulting from it, (by definition) takes place in Time, which immediately results in 4 complications for the non-critical editor and text-encoder:

1. Its beginning and end may be hard to determine and its internal composition difficult to define (document structure vs. unit of writing): authors frequently interrupt writing, leave sentences unfinished and so on.
2. Manuscripts frequently contain items such as scriptorial pauses which have immense importance in the analysis of the genesis of a text.
3. Even non-verbal elements such as sketches, drawings, or doodles may be regarded as forming a component of the writing process for some analytical purposes.
4. Below the level of the chronological act of writing, manuscripts may be segmented into units defined by thematic, syntactic, stylistic, etc. phenomena; no clear agreement exists, however, even as to the appropriate names for such segments.

These four complexities are exactly what the *TEI Guidelines* consider being distinctive features of speech. In chapter 11 on “Transcription of Speech” we read:

Unlike a written text, a speech event takes place in time. Its beginning and end may be hard to determine and its internal composition difficult to define. Most researchers agree that the utterances or turns of individual speakers form an important structural component in most kinds of speech, but these are rarely as well-behaved (in the structural sense) as paragraphs or other analogous units in written texts: speakers frequently interrupt each other, use gestures as well as words, leave remarks unfinished and so on. Speech itself, though it may be represented as words, frequently contains items such as vocalized pauses which, although only semi-lexical, have immense importance in the analysis of spoken text. Even non-vocal elements such as gestures may be regarded as forming a component of spoken text for some analytic purposes. Below the level of the individual utterance, speech may be segmented into units defined by phonological, prosodic, or syntactic phenomena; no clear agreement exists, however, even as to appropriate names for such segments. (258)

If we consider, as I propose, any holograph witness as a filtered materialization of an internal creative process (thinking) which can be roughly compared to an internal monologue or

dialogue, we may have a basis on which to build a methodology for the transcription of modern manuscript material. By combining the TEI DTD-subsets for the transcription of primary sources, the encoding of the critical apparatus and the transcription of speech, we could try to transcribe a manuscript and analyse it with tools for the manipulation of corpora of spoken language. It is interesting in this respect to observe how *critique génétique* describes the authorial interventions as deletions, additions, *Sofortkerretur* or *currente calamo*, substitutions, and displacements in terms of material or intellectual *gestures* as if it were kinesic (non-verbal, non-lexical) phenomena.

This approach does not do away with the essential problem of non-nesting information which is an inescapable fact of textual life, and in transcribing a holograph text even results from a one-way analysis.<sup>31</sup>

#### 4. Conclusion

Rather than focusing the debate on the possibilities of electronic scholarly editions and the advantages over the print edition in terms of usability—as has been too often the case—I'd rather see editorial theorists concentrate on the possibilities of the text which can be discovered by studying manuscript material very closely. Creating a non-critical edition/transcription of such a text with the use of text-encoding is the closest kind of reading one can endeavour. She who could reveal the writing out of the written work, discover the traces of the genesis of the text in its internal dynamics, could say something sensible about the text and the intentions of the author. But modern manuscripts are in their complexity unwilling to obey to the simple conventional ontologies of text and systems of text-encoding. Paraphrasing Louis Hay, one could say that “The principal merit of manuscripts is that they demonstrate the limitations and possibilities of text encoding. From the outset there are important material limitations: it is impossible to study a nonexistent manuscript.”<sup>32</sup> Paradoxically, the existent and extant manuscripts generate, by their resistance to current systems of text-encoding, new ontologies of the text and new approaches towards text-encoding. Likewise, current systems of text-encoding lay bare new possibilities by their limitations to cater for such manuscripts.



**Works cited**

- Barnard, D.T., R. Hayter, M. Karababa, G. Logan, and J. McFadden. "SGML-Based Markup for Literary Texts: Two Problems and Some Solutions." *Computers and the Humanities*, 22 (1988), 265-276.
- Barnard, D.T., L. Burnard, J.-P. Gaspart, L.A. Price, C.M. Sperberg-McQueen, and G.B. Varile. "Hierarchical Encoding of Text: Technical Problems and SGML solutions." *Computers and the Humanities*, 29 (1995), 211-231.
- Brians, Paul. "CD-ROMs in libraries & at home." Online posting. 7 Feb. 2000. *Humanist* 13.0380. 13 Jan. 2004.  
<[http://lists.village.virginia.edu/lists\\_archive/Humanist/v13/0374.html](http://lists.village.virginia.edu/lists_archive/Humanist/v13/0374.html)>.
- DeRose, Steven J. & Andries van Dam. "Document structure and markup in the FRESS hypertext system." *Markup Languages: Theory & Practice*, 1.1 (1999): 7-32.
- De Smedt, Marcel & Edward Vanhoutte. *Stijn Streuvels. De teleurgang van den Waterhoek. Elektronisch-kritische editie/electronic-critical edition*. Amsterdam: Amsterdam University Press, Gent: KANTL, 2000.
- & —. "The Best of Three Worlds: Eclecticism in Editorial Theory. The Electronic Edition of Stijn Streuvels' >>De teleurgang van den Waterhoek<<." *Sichtungen. Archiv-Bibliothek-Literaturwissenschaft*, 2001/2002 (4/5): 300-311. [Preprint available on <<http://www.onb.ac.at/sichtungen/berichte/smedt-m-1a.html>>].
- Driscoll, M.J. "Encoding Old Norse/Icelandic Primary Sources using TEI-Conformant SGML." *Literary and Linguistic Computing* 15.1 (2000): 81-91.
- Falconer, Graham. "Genetic Criticism." *Comparative Literature* 45.1 (1993): 1-21.

- Goldfarb, Charles F. *The SGML Handbook*. Oxford: OUP, 1995 (1990).
- Greetham, David. *Textual Scholarship. An Introduction*. New York/London: Garland, 1994.
- Grésillon, Almuth. *Eléments de critique génétique. Lire les manuscrits modernes*. Paris: Presses Universitaires de Paris, 1994.
- Goulet, Alain. "Devising and organizing a genetic edition of a literary work." *Digital Evidence. Selected Papers from DRH 2000, Digital Resources for the Humanities Conference, University of Sheffield, September 2000*. Ed. Michael Fraser, Nigel Williamson, and Marilyn Deegan. (Office for Humanities Communication Publication 14). London: Office for Humanities Communication, 2001. 179-189.
- Hay, Louis. "Does Text Exist?" *Studies in Bibliography* 41 (1988): 64-76.
- Hockey, Susan. *Electronic Texts in the Humanities. Principles and Practice*. Oxford: Oxford University Press, 2000.
- Kirschenbaum, Matthew G. *Image Based Humanities Computing*, thematic issue of *Computers and the Humanities* 36.1 (2002).
- Kline, Mary-Jo. *A Guide to Documentary Editing. Second Edition*. Baltimore and London: The Johns Hopkins University Press, 1998.
- Lavagnino, John. "Electronic Scholarly Editions." *Humanities Computing News. Newsletter of the Centre for Computing in the Humanities at King's College London*, February 2001. <<http://www.kcl.ac.uk/humanities/cch/hcn/hcn-2001-02.html#Vanhoutte>>.
- Lecolinet, Eric, Laurent Robert and François Role. "Text-image Coupling for Editing Literary Sources." *Computers and the Humanities*, 36.1 (2002): 49-73.

- Lernout, Geert. "‘Critique génétique’ und Philologie". *Text und Edition. Positionen und Perspektiven*. Ed. Nutt-Kofoth, Rüdiger, Bodo Plachta, H.T.M. van Vliet, and Hermann Zwerschina. Berlin: Erich Schmidt, 2000. 121-142. (trans. TEXT 14 [2002]: 53-75).
- McGann, Jerome. "The Rossetti Archive and Image-Based Electronic Editing." *The Literary Text in the Digital Age*. Ed. Richard J. Finneran. Ann Arbor: The University of Michigan Press, 1996. 145-183.
- . *Radiant Textuality: literature after the world wide web*. New York: Palgrave, 2001.
- Ore, Espen S. "Elektronisk publisering: forskjellige utgaveformer og forholdet til grunntekst(er) of endelig(e) tekst(er)." *Vid Texternas Vägskäl. Textkritiska uppsatser*. Eds. Lars Burman, and Barbro Ståhle Sjönell. Stockholm: Svenska Vitterhetsamfundet. 138-144.
- Ott, W. "Computers and Textual Editing." *Computers and Written Texts*. Ed. C.S. Butler. Oxford: Blackwell, 1992. 205-226.
- Reiman, Donald H. *Romantic Texts and Contexts*. Columbia: University of Missouri Press, 1987.
- Renear, Allen. "Literal transcription - Can the text ontologist help?" *New Media and the Humanities: Research and Applications*. Ed. Domenico Fiormonte and Jonathan Usher. Oxford: Humanities Computing Unit, 2001. 23-30.
- Robinson, Peter. "Manuscript Politics." *The Politics of the Electronic Text*. Ed. Warren Chernaik, Caroline Davis, and Marilyn Deegan. (Office for Humanities Communication Publications Number 3). Oxford: Office for Humanities Communication, 1993. 9-15.
- Robinson, P. M. W. (ed.). *The Wife of Bath's Prologue on CD-ROM*. Cambridge: Cambridge University Press, 1996.
- Ryle, Gilbert. *The Concept of Mind*. London: Penguin Books, 2000.

Schneider, Ben Ross, Jr. *Travels in Computerland; or Incompatibilities and Interfaces. A Full and True Account of the Implementation of the London Stage Information Bank. Reading etc.:* Addison Wesley, 1974.

Shillingsburg, Peter L. *Scholarly Editing in the Computer Age. Theory and Practice.* Third Edition. Ann Arbor: The University of Michigan Press, 1996.

—. “Principles for Electronic Archives, Scholarly Editions, and Tutorials.” *The Literary Text in the Digital Age*. Ed. Richard J. Finneran. Ann Arbor: The University of Michigan Press, 1996. 23-35.

Sperberg-McQueen, C. M. and Lou Burnard (eds.) *TEI P4 Guidelines for Electronic Text Encoding and Interchange. XML-compatible edition.* Oxford, Providence, Charlottesville, and Bergen: The TEI Consortium, 2001. <<http://www.tei-c.org/P4X/>>

Streuvels, Stijn. “De teleurgang van den Waterhoek.” *De Gids* 91.1 (1927), 64-118; 91.2 (1927), 153-181; 91.3 (1927), 316-342; 91.4 (1927), 1-43; 91.5 (1927), 137-179; 91.6 (1927), 285-316; 91.7 (1927), 1-63.

—. *De teleurgang van den Waterhoek*. Brugge: Excelsior [1927]; and Amsterdam: Veen [1927].

—. *De teleurgang van den Waterhoek*. Tweede herziene druk. Amsterdam: Veen, 1939.

—. *De teleurgang van den Waterhoek. Tekstkritische editie door Marcel De Smedt en Edward Vanhoutte*. Antwerpen: Manteau, 1999.

Tanselle, G. Thomas. “Reproductions and Scholarship.” *Studies in Bibliography* 42 (1989): 25-54.

—. “The Varieties of Scholarly Editing.” *Scholarly Editing: A Guide to Research*. D.C. Greetham (ed.). New York : Modern Language Association of America, 1995. 9-32.

—. “Textual Criticism at the Millenium.” *Studies in Bibliography* 54 (2001): 1-80.

Vander Meulen, David L. and G. Thomas Tanselle. "A System of Manuscript Transcription."

*Studies in Bibliography*, 52 (1999): 201-212.

Vanhoutte, Edward. "Where is the editor? Resistance in the creation of an electronic critical edition."

*DRH 98. Selected papers from Digital resources for the Humanities 1998. University of Glasgow, September 1998.* Ed. Marilyn Deegan, Jean Anderson, and Harold Short. London: Office for Humanities Communication, 2000 (OHC Publication 12). 171-183. A variant and earlier version of this article has been published in *Human IT*, 1.1999 (1999): 197-214.

<<http://www.hb.se/bhs/ith/1-99/ev.htm>>.

—. "A Linkemic Approach to Textual Variation. Theory and Practice of the Electronic-Critical Edition of Stijn Streuvels' *De teleurgang van den Waterhoek*." *Human IT*, 1.2000 (2000):

103-138. <<http://www.hb.se/bhs/ith/1-00/ev.htm>>.

—. "Display or Argument? Markup and Visualization for Electronic Scholarly Editions." *Standards*

*und Methoden der Volltextdigitalisierung. Beiträge des Internationalen Kolloquiums an der Universität Trier, 8./9. Oktober 2001.* Ed. Thomas Burch, Johannes Fournier, Kurt Gärtner & Andrea Rapp. Mainz: Akademie der Wissenschaften und der Literatur / Stuttgart: Franz Steiner Verlag, 2003. 71-96.

Vanhoutte, Edward & Ron Van Den Branden. *DALF guidelines for the description and encoding of modern manuscript material. Version 1.0.* Gent: CTB, 2002-2003.

<<http://www.kantl.be/ctb/project/dalf/>>.

—. "Presentational and Representational Issues in Correspondence Reconstruction and Sorting."

*Literary and Linguistic Computing*, 19/1 (2004): 45-54.

—. "Describing, Transcribing, Encoding, and Editing Modern Correspondence Material: a Textbase

Approach." Peter Shillingsburg & Fred Unwalla (eds.), *Computing the edition*. Toronto: Toronto University Press. 2004. (forthcoming).

Van Hulle, Dirk. *Textual Awareness. A Genetic Approach to the Late Works of James Joyce, Marcel Proust, and Thomas Mann.* PhD thesis. Antwerp: University of Antwerp (UIA), 1999.

Williams, Tony. *L'histoire en question. L'avant-texte du premier chapitre de la troisième partie de L'Education Sentimentale*. <<http://www.hull.ac.uk/hitm/>>.

Williams, Tony & Alan Brunt. "Hypertext and genetic studies: Flaubert's *L'Éducation sentimentale*." *Digital Evidence. Selected Papers from DRH 2000, Digital Resources for the Humanities Conference, University of Sheffield, September 2000*. Ed. Michael Fraser, Nigel Williamson, and Marilyn Deegan. (Office for Humanities Communication Publication 14). London: Office for Humanities Communication, 2001. 191-203.

## Notes

1. This essay is for Alois Pichler, director of the Wittgenstein Archives, whom I'd like to thank for the many fruitful discussions on textual genetics and transcription procedures, and for Heli Jakobson who managed to drag me from behind my desk every now and again during my stay in Norway. Research for this article has been made possible by a European Research Grant (5th Framework Programme Improving the Human Research Potential and the Socio-economic Base: Access to Research Infrastructures (ARI)) for research at the Wittgenstein Archives at the University of Bergen (Norway, 01/06/2002-07/07/2002).
2. "The Typopotamus called 'electronic edition'. Notes towards a definition and typology of electronic scholarly editions." King's College London, 8 december 2000.
3. Elsewhere I have suggested a model for electronic scholarly editing that unlinks the *Archival Function* from the *Museum Function*. "By *Archival Function* I mean the preservation of the literary artifact in its historical form and the historical-critical research of a literary work. *Museum Function* I define as the presentation by an editor of the physical appearance and/or the contents of the literary artifact in a documentary, aesthetic, sociological, authorial or bibliographical contextualization, intended for a specific public and published in a specific form and lay-out. The *digital archive* should be the place for the first function, showing a relative objectivity, or a documented subjectivity in its internal organization and encoding. The *Museum Function* should work in an *edition* — disregarding its external form — displaying the explicit and expressed subjectivity and the formal orientation of the editor. The relationship between these two functions is hierarchical: there is no *Museum Function* without an *Archival Function* and an edition should always be based on a digital archive." (Vanhoutte, *Resistance* 176).
4. See my and Ron Van den Branden's "Describing, Transcribing, Encoding, and Editing Modern Correspondence Material: A Textbase Approach" for a discussion on three ways to produce a scholarly edition with the use of text encoding, namely 1. Digitizing an existing print edition; 2. Creating an electronic edition e.g. by recording some or all of the known variations among different witnesses to the text in a critical apparatus of variants; and 3. Generating electronic editions from encoded transcriptions of the documentary source material.

5. For instance Jerome McGann's *Rossetti Hypermedia Archive* and Peter Robinson's *The Wife of Bath's Prologue on CD-ROM*.

6. "A *place* should be any piece of information, or at least any that exists in a stable or recoverable form." (DeRose & van Dam 9).

7. *De teleurgang van den Waterhoek* tells the story about the resistance of a rural community in a hamlet called de Waterhoek to the intrusion of industrial technology i.e. the building of a bridge, and the consequences this new connection to the outside world brings forth. The plot of the novel is complicated by the passionate relationship between the frank village girl Mira and Maurice, the engineer supervising the works. The text of the first print edition of the novel is 297 pages long, and counts 117,800 words.

8. Stijn Streuvels is a pseudonym for Frank Lateur.

9. These omissions were later addressed in a series of essays. See e.g. Vanhoutte, *Linkemic Approach*, De Smedt & Vanhoutte, *Eclecticism*, and Vanhoutte, *Argument*. The introductory material to the edition does give an account, however, of the textual genesis and the history of the text, descriptions of its physical forms and physical appearance, the rationale for editorial decisions such as emendations, the treatment of spelling, punctuation etc. A separate technical documentation outlines the markup strategies, the digitization process, the hardware and software used, and the creation of the electronic product.

10. For this purpose I will make extensive use of the aforementioned essays (see note 9).

11. The editorial project ran from 1998-1999 and was funded by the Royal Academy of Dutch Language and Literature, Gent-Belgium (Koninklijke Academie voor Nederlandse Taal- en Letterkunde <<http://www.kantl.be>>). The project was supervised by Prof. dr. Marcel De Smedt, who also functioned as a co-editor. Together with the "electronic-critical edition" on CD-ROM, a text-critical reading edition in bookform was published as a spin-off product of the electronic project (Antwerpen: Manteau, 1999).

12. together with a glossary, an introductory article on the genesis of the novel, a description of the transmission history of the text including all of the documentary sources and existing editions, an account of the principles underlying the constitution of the reading text (spelling, punctuation, corrections) with a list of corrections and end-of line hyphenations, an account of the principles underling the creation of the glossary, and a couple of facsimiles from the several documentary sources involved in the research.

13. Cf. Reiman esp. Chapter 10: 'Versioning': The Presentation of Multiple Texts, 167-180.

14. After an incubation period of twenty-some years, Streuvels started writing his novel and published it in installments in the Dutch literary journal *De Gids* in 1927. At the end of the same year it was published in bookform both in Flanders and the Netherlands by two different publishers using the same print. For the second print edition (1939) the text had been drastically reduced by 26.6% because the publisher requested a shorter and hence a cheaper version of the book. The author choose to cross out, amongst others, those passages that were

of a too explicit and erotic nature, and against which many catholic critics had fulminated in their reviews of the first print edition. At the same time, Streuvels changed the end of the novel. In the first edition the protagonist couple is divorced, but the author suggests that the divorced couple can eventually reunite, and there is an allusion to an unborn child. In the second drastically revised edition, the divorce is definitive and conclusive. Up to the publication of the 13<sup>th</sup> print edition of the novel in 1987, this drastically revised second edition had been the basis for 11 successful reprints.

15. Whereas the second print edition only retains 73.4% of the text of the first print edition, more emendations had to be made (93 and 73 respectively).

16. 76 in total.

17. These examples could respectively be paraphrased as `<CORR SIC="kiten" RESP="EV">kitten</CORR>` and `<SIC CORR="kitten">kiten</SIC>`.

18. The electronic-critical edition of *De teleurgang van den Waterhoek* is an auto-executable application which launches itself on inserting the CD-Rom in the CD-drive, and comes with the MultiDoc Pro CD Browser software. No programmes need to be installed on the hard drive of the computer (PC only) on which the edition is consulted. The electronic edition is an autonomous closed package without links to the internet. This meets the requirements Paul Brians (Washington State University) voiced in a discussion on the Humanist-list about CD-Roms in libraries and at home: “CD-ROMs at the least should be self-contained, and not require that files be installed on hard drives or permanent links be available to the Internet.”

19. For the transcription and markup of the correspondence material, I developed a project-specific StreuLet DTD which allows the encoding of letter-specific elements such as the existence of the envelope and envelope information such as postmark, place of posting, sender, sender’s address, recipient, recipient’s address, etc. This work has developed further into the *DALF guidelines for the description and encoding of modern correspondence material* and the DALF DTD. DALF is an acronym for “Digital Archive of Letters in Flanders” and focuses on correspondence by Flemish authors and composers from the 19<sup>th</sup> and 20<sup>th</sup> century. It is envisioned as a growing textbase of correspondence material which can generate different products for both academia and a wider audience, and thus provide a tool for diverse research disciplines ranging from literary criticism to historical, diachronic, synchronic, and sociolinguistic research. The input of this textbase will consist of the materials produced in separate electronic edition projects. See the papers by Vanhoutte & Van den Branden and the DALF website <http://www.kantl.be/ctb/project/dalf/>.

20. The choice of hardware and software, and the parameters decided on when batch converting a TIFF file to a lossy format such as Jpeg (e.g. the application of an Unsharp mask filter) are non-objective moments in the digitization process and highly influence the eventual result.

21. Annex C of ISO 8879 introduces the optional CONCUR feature (not available in XML) which “supports multiple concurrent structural views in addition to the abstract view. It allows

the user to associate element, entity, and notation declarations with particular document type names, via multiple document type declarations.” (Golfarb 89).

22. Suggested further reading on overlap-related problems: D.T. Barnard, R. Hayter, M. Karababa, G. Logan, and J. McFadden: SGML-Based Markup for Literary Texts: Two Problems and Some Solutions. In: *Computers and the Humanities*, 22 (1988), 265-276; D.T. Barnard, L. Burnard, J.-P. Gaspard, L.A. Price, C.M. Sperberg-McQueen, and G.B. Varile: Hierarchical Encoding of Text: Technical Problems and SGML solutions. In: *Computers and the Humanities*, 29 (1995), 211-231; Steven J. DeRose, D.G. Durand, E. Mylonas, and A. Renear: What is Text, Really? In: *Journal of Computing in Higher Education*, 1 (1990), 3-26; David Durand, Elli Mylonas, and Steve DeRose: What should markup really be? Applying theories of text to the design of markup systems. ALLC/ACH’96 Joint Conference of the ALLC and ACH, Bergen 1996; C. Huitfeldt: Multi-Dimensional Texts in a One Dimensional Medium. In: *Computers and the Humanities*, 28 (1995), 235-241; A. Renear, D. Durand, and E. Mylonas: Refining our Notion of What Text Really Is. In: *Research in Humanities Computing 4: Selected Papers from the 1992 ALLC/ACH Conference*. S. Hockey and N. Ide (eds.). Oxford, 1996, 263-280; C.M. Sperberg-McQueen, and Claus Huitfeldt: Concurrent Document Hierarchies in MECS and SGML. In: *Literary and Linguistic Computing*, 14 (1999), 29-42; C.M. Sperberg-McQueen, and Claus Huitfeldt: GODDAG: A Data Structure for Overlapping Hierarchies. <<http://www.hit.uib.no/claus/goddag.html>>; and chapter 31 “Multiple Hierarchies” of the *TEI Guidelines*.

23. The suggested methods are CONCUR, milestone elements, fragmentation of an element, virtual joints, and redundant encoding of information in multiple forms. Cf. Chapter 31 “Multiple Hierarchies” of the *TEI Guidelines*. The Rossetti Archive based in Virginia, and the Wittgenstein Archives at Bergen created their own encoding system—respectively RAD (*Rossetti Archive Document*), and MECS (*Multi Element Code System*)—out of dissatisfaction with the operationality of the options suggested by TEI. Cf. McGann, *Radiant* 88-97; and *Text Encoding at the Wittgenstein Archives* <<http://www.hit.uib.no/wab/1990-99/textencod.htm>>.

24. Cf. chapter 31 “Multiple Hierarchies” of the *TEI Guidelines*.

25. “manuscripts qui font partie d’une genèse textuelle attestée par plusieurs témoins successifs et qui manifestent le travail d’écriture d’un auteur.” (Grésillon 244).

26. Which I by all means consider a valuable and valid form of editing.

27. The absurd exponent of image based editing is the genetic edition of Flaubert’s *L’Éducation sentimentale* by Tony Williams and Allan Blunt (<<http://www.hull.ac.uk/hitm/>>). The diplomatic transcriptions were “initially prepared as Word (.doc) files, which cannot be readily converted into HTML text files, since ‘raw’ HTML code is unable to render the refined layout, freeform lines, and interlinear additions and other features of the transcriptions achieved through advanced use of modern word-processing software.” (Williams & Blunt 198). Therefore, the editors contend, “Diplomatic transcriptions need [...] to be stored in the hypertext package as image files.” (199) which they create by printing out the Word files and scanning them on a flatbed scanner. Hyperlinks from the transcription image file to

documentary notes are then established by means of image maps. But since hypertext is usually signalled by the browser by a different text colour, “the hotspots must be indicated (via words in red that do not corrupt the typographical integrity of the diplomatic rendering) in the word-processed transcription document before digitization.” (200)

28. Grésillon *Eléments* provides a good introduction to the methodology of critique génétique. A good introduction in English is Falconer, *Genetic Criticism*.

29. 86(3), May 1995.

30. “L’appareillage institutionnel et technique de la génétique textuelle ne saurait faire oublier que l’objet qu’elle se donne échappe presque par définition à la ‘science.’ Ce que scrute la génétique textuelle, c’est en effet un inobservable, un inobjectivable: l’origine même de l’oeuvre littéraire.” (Laurent Jenny, *Divagations généticiennes*, cited in Lernout 124).

31. See my *Display or Argument* for a further discussion of non-nesting information

32. The original citation: “The principal merit of manuscripts is that they demonstrate the limitations and possibilities of genetic criticism. From the outset there are important material limitations: it is impossible to study a nonexistent manuscript.” (Hay 68)