HAL open science

Encoding (inter)textual insertions in Latin "grammatical commentary"

Bruno Bureau, Christian Nicolas, Ariane Pinche

To cite this version:

Bruno Bureau, Christian Nicolas, Ariane Pinche. Encoding (inter)textual insertions in Latin "grammatical commentary". Journal of Data Mining and Digital Humanities, 2017, Special Issue on Computer-Aided Processing of Intertextuality in Ancient Languages, Special Issue on Computer-Aided Processing of Intertextuality in Ancient Languages, 10.46298/jdmdh.1392 . hal-01281266v2

HAL Id: hal-01281266

https://hal.science/hal-01281266v2

Submitted on 30 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Journal of Data Mining and Digital Humanities

Encoding (inter)textual insertions in Latin "grammatical commentary"

Bruno Bureau[1*], Christian Nicolas[2], Ariane Pinche[3]

1 Université Jean Moulin-Lyon 3, France

2 Université Jean Moulin-Lyon 3, France

3 Université Jean Moulin-Lyon 3, France

*Corresponding author: Bruno Bureau bruno.bureau@univ-lyon3.fr

Abstract
The ancient commentaries provide a large sample of quotations from classical or biblical texts for which Latin grammarians developed a complex system of insertion of quoted texts. The paper examines how to encode such passages (using XML Tei, and focuses on difficult cases, such as inaccurate quotations, or quotations of partly or wholly lost texts.

Keywords
Ancient commentary; quotation; allusion; fragmentary texts ; mentioned words ; Latin grammar ; XML Tei.

INTRODUCTION

"Grammatical commentary" is a very important but not the best-known part of Latin literature. [Goulet-Cazé and Dorandi, 2000], [Geerlings & Schulze, 2002] and [Geerlings & Schulze, 2004] provide recent general introductions to the genre and its various problems. Under this generic name, we find texts from various places and dates and using different methods. Some are clearly a collection of notes that teachers probably used as a basis for their lessons[1], some are more sophisticated productions that include a biographical, historical and sometimes technical erudite material, and were probably regarded as a kind of companion to Vergilius, Terence, Statius etc. [2] In some cases, commentaries provide complete essays on a work, with a real attention paid to the literary value of the commentary itself as a literary

[1]This material mostly includes the *scholia*, marginal or interlinear glosses, which each scholar could rework on his own, by including new material or omitting notes if he considered them of less value or interest. The creative process of such material is visible in multi-hand glossated manuscritps as for instance Paris BnF Lat. 18554 from the middle 9th century. In this copy of Arator's poem *Historia Apostolica* (from the 6th century), a scholar completed with his own notes an inherited corpus of glosses (that we know from other sources) among which he, or his scribe, selected only a small part.

[2]Such are the « great commentaries » of Porphyrio, Donatus, Servius, Lactantius Placidius and others, but also more problematic texts as the *adonationes in Lucanum*. In Christian area, Cassiodorus' work on the Psalms and a part of Jerome's commentaries can be ranked in this category. Among Christian commentaries, we set apart a series of texts that actually consist in a collection of homelies such as Augustine's treatises on John's gospel and *enarrationes in psalmos* and for this reason have not the same literary status.

production.[3] More than a literary genre, these texts witness the various attempts and methods of ancient Latin scholarship (about which see [Kaster, 1988]). Under the names of Servius, Donatus, Lactantius, what the manuscripts provide is actually the result of a long and complex tradition that, though it takes its origin in the real work of these famous scholars, is also the result of multiple reworkings throughout Antiquity and the Middle-Ages. These preliminary remarks (well known to scholars especially devoted to the study of these texts, but not familiar with the others) are of great importance for anyone who tries to understand and study the mechanisms of quotation, (and more generally textual insertion in these texts) and consequently faces in a digital edition the problem of encoding the very fragmentary and unordered textual material provided by these "texts".

As this paper will partly focus on Donatus' use of quotation, we shall begin with a few details concerning Donatus' textual transmission because we cannot properly understand Donatus' system of quotations if we do not pay due attention to the very special history of Donatus' text. On Donatus himself as a commentator, see [Jakobi, 1996], and on textual criticism issues [Reeve, 1978] and [Reeve, 1979]. Donatus was regarded by ancient and mediaeval scholars as one of the greatest if not the greatest Roman grammarian. He was known both for his treatises or *artes* and for at least two commentaries, one of Vergilius (now lost except for the Vergilius' life and the very beginning of Bucolic 1) and one of Terence (now extant but with a very high probability of complex reworking). [Holtz, 1984] and [Holtz, 2000] clearly demonstrated that Donatus' tradition of Terence commentary could be divided in three steps: 1-Donatus wrote a continuous commentary of Terence's plays, in a form that could be very similar to Cassiodorus' preserved commentaries on the Psalms or Tiberius Donatus' on Vergil. The commentary was a coherent text separated from the text of the comedy, but including quotations of Terence's plays either integrated in Donatus' phraseology, or used as catchwords to introduce a commentary on one or several lines, the *lemmas*. The reader had to use two books: Terence's text in one hand, and Donatus' book in the other. 2-During the early Middle-Ages, the commentary was reintegrated in Terence's text of the comedies, in order to avoid the difficulty of using at the same time two separate *codices*. The commentary was cut into pieces and placed near the commented words of the comedy as marginal or interlinear glosses, as it is usual in most of mediaeval manuscripts of classical Latin authors. 3-Probably during the 9th century these glosses, which were generally identified as a part of Donatus' commentary with such formulas as *Donatus dixit*, *sic Donatus*, were reunited in a continuous text. We can see the result in the oldest preserved manuscript of Donatus' commentary, Paris BnF 7920 (from the 11th century).

On the contrary, Cassiodorus' commentary of the Psalms has been preserved in a complete and quite stable form since it has been published probably between 560 and 580. The manuscript tradition is very similar to the traditions of literary texts and shows very few (if any) trace of reworking, except for Cassiodorus' own revisions (on this text, see now [Stoppacci, 2012]).

I DIFFERENT TYPES OF QUOTATIONS IN GRAMMATICAL COMMENTARIES.

[3]The most fascinating example is Tiberius Donatus' commentary on Vergil's *Aeneid*. The twelve books offer a complete reading of the poem that includes in an elaborated form narrative analyses, psychologic and dramatic notes, grammatical and stylistic issues. Cassiodorus' on Psalms is somewhat like a Christian attempt of the same kind of commentary, but its links to grammatical commentaries are far tighter.

Many scholars have been interested in ancient mechanisms of quotation, and two recent studies, [Darbo-Peschanski, 2005] and [Nicolas, 2006], give a general overview of the problems. We will examine in the following three very frequent cases in commentaries: the *lemma*, internal and external quotations and the mentioned words.

1.1 The *lemma*

As [Holtz, 1984] and [Holtz, 2000] already pointed it out, this transmission had very important consequences on one type of quotations, what we call the *lemma*, namely the part of Terence's play used to introduce the commentary. We do not know how exactly the system was in Donatus' first redaction, but we can hold for certain that what we now identify as lemmas is probably the result of a mediaeval reconstruction of the text. Generally, this *lemma* consists in a sequence of contiguous words that point to a single line, a group of lines or sometimes a whole scene. The length of the commentary assigned to a single *lemma* highly depends on the nature of the explanations provided by Donatus. Discontinuous quotations in the *lemmas* are quite exceptional and most of time point out a textual issue so that we can hold for very likely that the rule was to use continuous quotations in the *lemmas*. Every commentary, except perhaps Tiberius Donatus' on Vergil, is based on this system of quotation of the commented text. Donatus is so a very good pattern to test our encoding procedures.

1.2 Internal and external quotations

We name internal quotations quotations of other parts of the work, which were commented by the grammarian elsewhere in the corpus and are used as intratextual material, and intertextual quotations of other Greek and Latin works. For instance, in Servius commentary of *Aeneid* book 1, 234, either a quotation from book 1 (for instance 1, 542) or a quotation of book 11 will be considered as internal quotations. On the contrary Buc. 3, 22 will be treated as external quotation even though Servius has also written a commentary on Bucolics. External quotations in such commentaries are generally easily identifiable, but in some very interesting cases, we are dealing with more or less known fragments of more or less identified works[4]. As the grammarian most of the time gives no precise indication (as we usually do in modern citations), and uses sentences such as *ut ait quidam*, *Cicero in rhetorica* or *apud ueteres patres*, it is sometimes really difficult and near impossible to find out where he has taken the fragment of text that he inserted in his note.

In Donatus' specific case, the quotations from Terence's *Self-Tormentor* are problematic, because we do not know with certainty if the grammarian wrote this commentary which in this case is now lost or if (for what reason, we do not know) he never completed the commentary of the six plays. We chose to treat them as internal quotations, according to the fact that it is most likely that Donatus wrote this part of the commentary and that it was lost.

1.3 Accuracy in quotations

A very disturbing problem is set by the very numerous cases where the quotation in the commentary does not fit exactly modern editions (or even their apparatus) of the quoted text.

Some cases are very easy to explain: for instance, Cassiodorus read the Psalter in a Latin translation used in Italy in the 6th century and which is slightly different from the Vulgate's Gallic psalter, and very different from Jerome's new translation based on Hebraic material [Gribomont, 1986]. In those cases, accuracy is not really a problem, except when (what he sometimes does) Cassiodorus quotes in internal quotations a text that is different from the text

[4] See below for a very interesting case of false "accurate quotation".

that he will comment in its due place. We may suspect that the exegete quoted from memory and had no time to check every quotation in his huge work[5].
Despite these easy cases, accuracy of quotations in commentary is a real and difficult problem.

1.3.1 What is an accurate quotation?

A very simple way to define accuracy would be to state that a quotation is accurate when it perfectly reproduces the quoted text. In ancient literature, therefore such a definition is nonsensical. We perfectly know that there were very important textual differences in the numerous copies of the same text, and the more known and read the text was, the more numerous the differences were. Besides obviously erroneous readings that grammarians pointed out in the copies that they used, there was a kind of medium state containing variant readings, that were metrically and stylistically correct, and among whom it was very difficult to determine which the original one was. There were also variant readings coming from the writer himself, when he had tried several versions of the same line or sentence or corrected for a new edition a text whose old version was still at hand. We know from late antique examples that this was a common practice, and we know for instance from epistolography and even though no different versions had been preserved that classical writers did exactly the same (see for example [Gurd, 2007]). [Stoppacci 2009] gives a very new and suggestive account of such reworking in Cassiodorus' text, but Prudentius' poems were probably reworked by their author. On this topic see [Cunningham, 1968] and [Gnilka, 2010].

1.3.2 What means inaccurate quotation in the commentaries?

In this context, apparent inaccuracy may often hide real textual variants. Before considering that a quotation is inaccurate, and eventually before correcting it, we must pay due attention to the critical apparatus of modern editions of the quoted text, to determine if the variant that we read in the commentary is a part of the textual tradition of the text or not. Even when it seems that no preserved manuscript reads the same text as the commentator, we may often suppose that this reading (except when it is obviously an erroneous one) might have been read in a now lost copy. In fact, inaccuracy is, if we regard it with these due precautions, a relatively rare phenomenon in the preserved commentaries. We must keep in mind that grammarians were (as said [Kaster, 1988]) "the guardians of language" and probably attached great importance to accuracy regarding the classical texts. In Christian area, the commentators of the sacred text were probably even more accurate with what they held for God's own words. These remarks explain why, in our encoding process, we tend to limit to obvious mistakes the cases where we suppose (and encode) an inaccurate quotation.
This practice of quotation is often mistaken for another practice, the allusion, which may sometimes leads us to regard as an inaccurate quotation words that in fact are no quotation, but a simple allusion. This case is especially frequent when commentators want to focus on

[5] A different case is Cass. *ps.*, 1, 1 where a quotation from Zechariah is attributed to Isaiah: *Nec dubites quod uirum appellat Dominum Saluatorem, de quo et Isaias propheta dicit: "ecce uir Oriens est nomen eius"*. In this special case, the quotation itself is right but the citation is erroneous. Such a mistake can be explained by Cassiodorus' source at this point. While explaining Is. 41, 25, Jerome wrote *et in alio loco: ecce uir oriens nomen eius.* Cassiodorus may have understood "elsewhere in his book" instead of "in another book" and falsely attributed the quotation to Isaiah. Such erroneous citations are not exceptionnal.

the idea and not to analyze the form of the "quoted" text. In these cases, we often face "quotations" in which, if the main words of the original are preserved, the writer often changes what he considers as details or elements of less interest. So we had to define what was quotation and what allusion and adapt our encoding to both cases.

II ENCODING BASIC QUOTATIONS

Under "basic quotations", we mean quotations in which the quoted text is well known, accurately quoted and without any reformulation such as[6]:

> *inclementivs pro 'inclementer', ut 'iam senior, sed cruda deo uiridisque senectus'* (Don. *Eun.* 4, 3).

Our encoding system uses the different values of <quote> and <cit> tags and eventually @type and @subtype to give more precisions. The quotation above may be encoded as

```
<cit type="Latin poetry" subtype="epic">
        <quote>
                iam senior, sed cruda deo uiridisque senectus
        </quote>
</cit>
```

The @type and @subtype here are defined by the use of each commentator. For Donatus, it may be useful to structure his quoted material in "prose" vs. "poetry", "Greek" vs. "Latin" and to characterize the type of literary productions as "epic" or "elegiac" and so on. With this encoding, it becomes obvious that we consider XML Tei documents as databases whose entries are defined according to the objects that we regard as pertinent data.

This kind of encoding is reused in a different way for Cassiodorus commentary on Psalms, for instance in this passage (Cassiod. *ps. Praef. epist.*) :

```
<cit type="biblical" subtype="OT">
        <quote>
          fons signatus, paradisus plenus omnium pomorum
        </quote>
</cit>
```

In this case, @type and @subtype are defined in a different way, according to what seems important to put in our database, namely here the difference between "biblical", "patristic" and "secular" quotations, and among "biblical" quotation the difference between Old Testament and New Testament quotations.

The point here was to make a clear difference between lemmas and quotations. We consider that a lemma is a quotation because it inserts in the commentary something that is neither a commentary nor a citation, because there is no choice left to the commentator to choose the work from which he will take the quotation. In the commentary of the *Girl from Andros*, for instance, Donatus must exclusively use as *lemmas* words coming from this play. Our encoding reflects this difference by using only the <quote> tag, with @type="lemma", as in the following example from Donatus' *Andr.* 5 :

```
<quote type="lemma">
        opera abutitur
</quote>.
```

Another problem concerns the citations, and the best way to produce readable citations that may also be used for other databases or other projects. Until now, we chose to use standardized citation patterns, namely the TLL abbreviation corpus for Latin texts and Liddell-Scott system for Greek material. These elements are, as usual in TEI documents, in a <ref> tag with @cRef, as in the passage quoted below from Donatus' *Eun. 4, 3* :

```
<cit type="Latin poetry" subtype="epic">
        <quote>
                iam senior, sed cruda deo uiridisque senectus
        </quote>
        <ref cRef="Verg. Aen. 6, 304 "/>
</cit>
```

[6] We quote Donatus according to [Wessner, 1902] with our emendations [Bureau and Nicolas, 2012]

This system provides in our database a systematized corpus of citations, in a form that is familiar to scholars, but is not really sufficient for other uses of our data, especially if we want to cross our data with other projects or allow the reader to access directly the full text of the quoted work. As the target and cRef attributes are mutually exclusive, according to TEI guidelines, a solution could be the use of URIs when they exist. The example below could be tagged:

```
<cit type="latin poetry" subtype="epic">
        <quote>
                iam senior, sed cruda deo uiridisque senectus
        </quote>
        <ref target="http://data.perseus.org/citations/urn:cts:latinLit:phi0690.phi003.perseus-lat1:6.295-6.336 "/>
</cit>
```

to point to the Perseus' Vergil. In fact, the help provided by this system is still deceptive. Many late Latin texts are not available as targets through URI system, and in many (if not most) cases we should have to use our @cRef system. It is the reason why for the moment and in the present state of our corpus, we decided to use only the @cRef system.

A last kind of unproblematic quotation consists in "mentioned elements".[7] These elements are parts of speech not used for their meaning but for their grammatical value, such as for example "*Donatus* is a proper name" as compared to "Donatus was a grammarian". On this mentioned elements, see [Nicolas, 2005] and [Nicolas, 2007]. As expected, we used the TEI element "mentioned" including @lang for mentioned Greek words (Latin is regarded as default language).[8] Nevertheless, even with these elements the encoding of which was apparently easy, we encountered some problems that involve the meaning and the purpose of our encoding system of quotations.

III PROBLEMATIC QUOTATIONS AND THEIR ENCODING METHOD: FRAGMENTARY, INACCURATE AND REFORMULATED QUOTATIONS

When we tried to encode problematic quotations, we encountered three main problems. 1-Some obviously "mentioned" terms are quoted in an apparently inaccurate form that is often but not always involved by the syntactic structure of the passage that includes the quotation. 2-Some quotations are obviously accurate but include one or more variant(s) in regard of modern editions of the quoted text. 3-Sometimes, the quoted text is partly or totally unknown to us, mainly because it comes from a (partially or totally) lost work. For those three cases, we had to find proper encoding methods that could give the finest possible accuracy in the description and treatment of ecdotic problems.

3.1 "Inaccuracy" in mentioned words

[7] Mentioned terms are unproblematic when their character of mentioned part of speech is obvious. If we say "*horse* is the name of an animal", the word "horse" is undoubtedly mentioned. However, not all cases are so easy. For complex cases, see below § 3.

[8] The problem of "default language" is in itself interesting. We admitted that "default language" was the language of the largest part of the text, Latin for a Latin text, Greek for a Greek one. However, how far we can extend such a choice is not very clear to us. If we consider that Latin is our default language, how shall we represent Latin parts of speech? According to Latin rules (in which, for instance, nouns and adjectives are the same part of speech) or according to modern rules? Moreover, if we choose modern rules, from what language shall we take them? This question becomes a very important point when we intend to treebank our corpus.

While preparing an encoding schema for the critical digital edition of Donatus' commentary, we realized that we had not paid due attention to a very special case of mentioned elements that we could name "partially mentioned words". In some cases, Donatus uses a mentioned word in a form that includes this word in a phrase that we cannot understand if we just read the word as a grammatical object, a merely mentioned word. In these cases, the meaning of the word is implied in the grammatical or lexical analysis. Let us consider one example about a (fictitious) *lemma* like "*te diligo immo te amo*" (I do not just like you; I do love you). If Donatus writes, "*amare* is the word for love, *diligere* the word for friendship", there is no problem; but if he writes: "*amamus* the beloved persons, *diligimus* our friends", we cannot treat *amamus* and *diligimus* like purely mentioned words (because the meaning of the verbs is important if we want to understand the phrase). Neither can we regard these words as quotations because their form (1st person plural) is quite different from the form used by the poet (1st person singular). Making a decision in either direction (mentioned or quoted word) would be a mistake, because the interest of such a place is precisely that the writer maintains an ambiguous position. A solution can be provided by the attribute @ana (@type is not allowed in <mentioned> tag) which gives the exact nature of the mentioned word. So we can write in a different way:

```
<mentioned>amare</mentioned>
is the word for love,
<mentioned>diligere</mentioned>
the word for friendship
```

and:

```
<mentioned ana="included">amamus</mentioned>
the beloved persons,
<mentioned ana="included">diligimus</mentioned>
our friends
```

where "included" means that the mentioned word is syntactically linked to the rest of the sentence.

Latin etymologies provide a special case, when the word the etymology of which is in question, is related to its etymon by *a* or *ab* that implies both the use of ablative and sometimes (especially in verbs) a change in the form of the verb from infinitive to gerundive. Is a sentence as *amor dicitur ab amando* (the word "love" comes from the verb "to love") a case of free or linked mentioned term? For the moment, we consider that in this special case the mentioned term is free because the changes are not the result of a stylistic choice by the writer as it was in *amamus*, but only the consequence of the syntactic rules applying to the way of introducing the etymology. This choice however may be questioned if we find examples that are more ambiguous.

3.2 Encoding textual variants in an "accurate" quotation

Cassiodorus' commentary on the psalter provides an interesting set of textual variants in biblical text that perfectly illustrates the problem of textual variants in quoted places of extant works. As we know, Cassiodorus uses an Italian Latin Bible, which is sometimes quite different from the Vulgate and is an interesting document on biblical text in the mid-6th century Italy. For our purpose, which is not a new critical edition neither of Cassiodorus nor of his biblical text, we decided to treat the places in which Cassiodorus' text is different from the Vulgate as textual variants of the same text. We are fully aware that this editorial choice is partly an oversimplification of the complex tradition of the Latin Bible, but we hold for important to individuate these passages even if we cannot deal with the whole history of their

transmission. An example for our encoding method in such places could be the following (Cass. *ps*. *Praef*. 15[9]):

```
Sic enim in euangelio de Domini Christi praedicatione dictum est:
    <cit type="biblical" subtype="NT">
      <quote>
        <app>
          <lem wit="#CassiodPsalm">loquebatur autem</lem>
          <rdg wit="#Vulg">erat enim docens eos</rdg>
        </app>
        sicut potestatem habens, non sicut scribae eorum et pharisaei
      </quote>
      <ref cRef="mt. 7, 29"/>
    </cit>.
```

We use apparatus tags to indicate the variant without any further precision on the Vulgate tradition. The point here is not to give a detailed apparatus of this quotation of the Gospels but to show a possible place for a further enquiry on the Italian version of Matthew. But, in fact, this quotation is more problematic than it seems at first glance. Where does in fact the quotation begin? Is the segment *loquebatur autem* (but he was speaking) a part of the quotation or a reformulation of the sacred text which could be *erat enim docens* (for he was teaching)? No other patristic text gives the segment *loquebatur autem*, and so it could be inferred that the first segment is rather a reformulation of the text than a real variant and so our encoding may be:

```
<cit type="biblical" subtype="NT">
        <quote>
          <app>
            <lem wit="#CassiodPsalm" type="reformulation">
                    loquebatur autem
            </lem>
            <rdg wit="#Vulg">erat enim docens eos</rdg>
          </app>
          sicut potestatem habens, non sicut scribae eorum et pharisaei
        </quote>
        <ref cRef="mt. 7, 29"/>
  </cit>
```

The @type may here be used not to indicate what kind of variant this reading is but where the variant may originate. Other possible values could be "translation" or "allusion" and default value would be what we generally regard as a textual variant.

This apparently easy example shows how accurate the encoding work must be in quotations. As we will see below, the main issue here is to provide a reusable material by including in our XML text what is important for a further use of our data.

3.3 Encoding fragmentary quotations: how to deal with quotations of lost texts?

Another disturbing case is the case of fragmentary quotations. Everyone who works on Ancient commentaries is confronted with fragments of lost works, and we intended to treat these places exactly in the same way as the other external quotations. However, in the following case (*Pho*. 171), this easy way is quite inappropriate. Donatus indeed writes:

> *Sallustius libro secundo "et Poeni fere aduersus a. a. n. n. e. e. m. m."*
>
> Sallust in the second book of his *History* "and the Carthaginians against…"

and we don't know "against what?", because that is the only preserved quotation of this fragment. Usually when the manuscripts use abbreviated forms of quotation, by writing only the first letter of each word, we use the element <choice>, with <abbr> and <expan> elements for encoding the abbreviated form and the expanded one of the quotation. In this case, we can't. What was a well-known quotation for the grammarian is a partly lost fragment for us. In front of such cases, we have in fact two solutions: the simple one is to let the <expan>

[9] P. Stoppacci has not yet completed her new critical edition, which will greatly improve our knowledge of the textual transmission, so we use [(Cassiodorus, 1958)], which is not a critical edition, but provides a generally suitable text.

element void. The second one, that is more sophisticated, is to include in <expan> element a possible reconstruction of the lacking words (for example by using an edition of Sallust's fragments) and to use an attribute @resp, meaning "according to such or such modern edition". In the present case, the encoding could be:

```
<cit type="latin prose" subtype="historiography">
        <quote>
                et Poeni fere aduersus
                        <choice>
                                <abbr>a. a. n. n. e. e. m. m.</abbr>
                                <expan/>
                        </choice>
        </quote>
</cit>
```

In fact, the problem is much wider than the relatively rare case of otherwise unknown fragments. If Donatus' manuscripts write *ut Verg. a. u. q. c.*, can we really suppose that the grammarian was meaning *arma uirumque cano* and not for instance *arma uirumque canam*? In fact, we think that he was meaning the received text of the first line of the *Aeneid* just because this line is one of the most famous quotations in Latin poetry, but actually, we cannot be sure without any doubt. This fact implies that we cannot treat this kind of quotations exactly as fully developed ones, especially if we keep in mind that the text quoted by Donatus is often different from the text that is consensually received today. Such a place may be encoded as:

```
ut
<choice>
        <abbr>Verg.</abbr>
        <expan>Vergilius</expan>
</choice>
<cit type="Latin poetry" subtype="epic">
        <quote>
                <choice>
                        <abbr>a. u. q. c. </abbr>
                        <expan cert="high"> arma uirumque cano</expan>
                </choice>
        </quote>
</cit>
```

where @cert is used to underline the fact that the expanded form is very likely but not certain. On the contrary, the default value in the first choice means that the expanded form is regarded as undoubted.

To close this part of the paper, let us consider a very interesting case of an inaccurate, fragmentary and unknown quotation. In *Pho*. 2, Donatus' previous editor reads:

leuibus huic hamis c. a. t. d. h., u. d. et t. i. A

in which the abbreviated segment is identified as a quotation of *Aen*. 5, 259 and 5, 262, according to a Robert Estienne's conjecture from his 1529 edition. However, when we read the manuscripts, we can see that they consensually read a very different text:

leuius ter hinc anus conseruatam c a a/t id dedit h u d e decus erat[10].

The differences are such that we can suspect the great French scholar of having here overused his (considerable) ability to make useful emendations. If we stand close to manuscripts this text is not a Virgilian quotation but a new fragment of the Latin poet Laevius. Estienne misunderstood the word *leuius* and regarded it as an adverb, because he had probably no idea of who the poet Laevius was. If we admit that the first word is a proper name, the following words are not in so desperate condition that we could not rebuild something like a Latin sentence:

[10]Of few interest for our purpose here are the slight alterations of this text in manuscripts in which we find *annus* (year) instead of *anus* (old woman) and *decusserat* (he/she has stroken down) instead of *decus erat* (it was an honour); such graphic variants with word confusion or misspelling are very common.

Laeuius: ter hinc anus conserua tam ca<te> id dedit † h u † dedecus erat

and the poet Laevius: 'thrice the old woman who was his companion in slavery so craftily gave it, <and that> was a shame'[11].

Whatever be its interest for ecdotic reasons, from the point of view of the encoding system, such a place is very problematic. We must indeed indicate that this place is a quotation, but coming from a correction from the editors (us), which is in fact a simple return to the text of the manuscripts and which includes a partly meaningless text. According to what we previously said, these words may be encoded as follows:

```
<cit type="latin poetry" subtype="unknown">
  <quote>
    <app>
        <rdg wit="#mss">
         ter hinc anus
                <choice>
                        <sic>conseruatam</sic>
                        <corr resp="HD" cert="high">conserua tam </corr>
                </choice>
                <choice>
                        <sic>caa/t</sic>
                        <corr resp="HD" cert="medium"> cate</corr>
                </choice>
        id dedit
                <choice>
                        <sic>h u</sic>
                        <corr resp="HD" cert="low"> hoc</corr>
                </choice>
        dedecus erat
      </rdg>
  </app>
 </quote>
</cit>
```

CONCLUSION

To conclude these remarks with a few words, it seems to us that this encoding process, which is complex and sometimes undoubtedly questionable, implies a new relation to the texts, and perhaps a new reflection on the ecdotic process in itself. XML Tei helps us to change a simple text into a structured database, in which each encoded element is ranked in a series of similar or near-similar ones. Clearly, our purpose is not to give final solutions to the various problems that we described in the preceding pages, but to set the problems on the table by identifying in each text of our corpus passages for which a methodological reflection is required.

As far as quotations are concerned, the mechanisms of intertextual insertions in ancient texts is a very complex and varied process that implies both linguistic and stylistic questions that the editor must keep in mind during the collation of manuscripts and the phase of textual choices. The quotations are for scribes (and consequently for editors) dangerous places where the limits of the quoted and the quoting texts are sometimes blurred. A major part of our

[11]Many other details support this interpretation. 1-The preceding commentary which Donatus intends to illustrate with this example, deals with redundant phrases as *homo seruus* (a slave man) instead of *seruus* (a slave). If we read like Estienne, the example is irrelevant, but in our interpretation *anus conserua* (an old woman companion of slavery) is a perfect example of this kind of phrases. 2-Donatus always quotes consecutive lines and this case would be the only one of such fragmentary quotation. 3-The Laevius' segment is metrically the beginning of iambic verse (most likely a trimeter): *tēr hīnc/ ănūs/ cōnsēr/uă tām/ ca<te> id dedit.* We do not know much about Laevius' poetry, but his use of iambic verses is certain.

encoding work is not to dispel this ambiguity, but on the contrary to show it and to give the most precise view of the incertainties and put these passages into series of similar and equally ambiguous ones. As a second work, when we change our role to the one of the editor, we use these series to compare the cases, and finally to choose what we consider as the best (or at least a satisfying) solution. The main advantages of the XML treatment of the quotations is precisely the building of such patterns for representing intertextuality that can help the editors in their work.

(Donat, 1902)

("Hyperdonat | Hyperdonatus - Editiones collectae antiquorum commentorum electronicae cum translatione, commentariis et adnotationibus criticis.," n.d.)

References

Bureau B., Nicolas C., Hyperdonat, édition électronique de commentaires anciens, Donat, *Commentaire des comédies de Térence*, 2012: http://hyperdonat.huma-num.fr/

Cassiodori, *Expositio psalmorum*, P. Stoppacci, ed. Firenze, Sismel : Edizioni del Galluzzo, 2012.

Cunningham M. P. The Problem of Interpolation in the Textual Tradition of Prudentius. *Transactions and Proceedings of the American Philological Association*, 1968: 99: 119-141.

Geerlings, W., Schulze, C. (eds.). *Der Kommentar in Antike und Mittelalter: Beiträge zu seiner Erforschung*. Leiden, 2002.

Geerlings, W., Schulze, C. (eds.). *Der Kommentar in Antike und Mittelalter: neue Beiträge zu seiner Erforschung*. Leiden, 2004.

Gnilka C. Versdubletten Zu Prudentius, Peristephanon 2,514. *Vigiliae Christianae*, 2010: 64, 5: 492-495.

Goulet-Cazé M.-O. *et al.* (eds), *Le commentaire : entre tradition et innovation, Actes du colloque international de l'Institut des traditions textuelles (Paris et Villejuif, 22-25 septembre 1999)*, Bibliothèque d'histoire de la philosophie, Paris, 2000.

Gribomont, J. Cassiodore et ses bibles latines. *Atti della settimana di studi su Flavio Magno Aurelio Cassiodoro (Cosenza-Squillace 19-24 settembre 1983)*, Mannelli Rubbettino Ed. 1986: 262–280.

Gurd S., Cicero and Editorial Revision, *ClAnt* 2007: 26, 1: 49-80.

Holtz L., Les manuscrits latins à gloses et à commentaires de l'Antiquité à l'époque carolingienne, *in* C. Questa, R. Raffaelli (eds), *Atti del Convegno internazionale «Il libro e il testo»: Urbino, 20-23 settembre 1982*, Pubblicazioni dell'Università di Urbino. Science umane 1, Urbino, 1984: 141-166.

Holtz L., Le rôle des commentaires d'auteurs classiques dans l'émergence d'une mise en page associant texte et commentaire (Moyen-Âge occidendal), *in* Goulet-Cazé M.-O. *et al.* (eds), *Le commentaire : entre tradition et innovation, Actes du colloque international de l'Institut des traditions textuelles (Paris et Villejuif, 22-25 septembre 1999)*, Bibliothèque d'histoire de la philosophie, Paris, 2000 : 101-118.

Jakobi R., *Die Kunst der Exegese im Terenzkommentar des Donat*, De Gruyter, Berlin, New York, 1996.

Kaster R. A. *Guardians of language: the grammarian and society in late antiquity*. Berkeley, 1988.

Nicolas C., Typologie de l'autonymie en latin, *in* Poli F., Vottéro G., *De Cyrène à Catherine : trois mille ans de Libyennes. Etudes grecques et latines offertes à Catherine Dobias-Lalou*, A.D.R.A., Nancy, 2005.

Nicolas C., Quel est le statut syntaxique des autonymes dans les énoncés latins ? L'exemple des tours "X ab Y", *in* Bodelot C., *Eléments "asyntaxiques" ou hors structure dans l'énoncé latin*, Presses de l'Université Blaise Pascal, Clermont-Ferrand, 2007.

Reeve M. D., The Textual Tradition of Donatus' Commentary on Terence, *Hermes*, 1978: 106, 4: 608-618.

Reeve M. D., The Textual Tradition of Donatus' Commentary on Terence, *Classical Philology*, 1979: 74, 4: 310-326.

Stoppacci, P. Stadi redazionali nella tradizione manoscritta dell'« Expositio psalmorum » di Cassiodoro. *Studi Medievali*, 2009: 3: 499–557.