Linguistic and Philological Variants in the Papyri: A Reconsideration in Light of the Digitization of the Greek Medical Papyri

Nicola Reggiani

It might be not so original to start with the traditional description of a variant as a deviation of a text from its archetype, but here exactly lies the similarity between linguistic and philological variants, on which the following pages will be focused. Both conceal the assumption that we need to emend a text in order to reach a virtual textual exactness with reference to one, single archetype, and in both cases the critical editor will print what he assumes to be the ‘correct’ form in the text, relegating the deviating ‘anomaly’ in the apparatus.

While a philological variant is usually defined after a comparison with another version of the same text, papyrus documents in most cases appear to be unique texts. They are, according to the terminology of textual criticism, ‘single witnesses’, and their ‘variants’ and ‘errors’ are usually intended as related not to an archetypical text, but to a standard reference language: Koine Greek. One of the most striking editorial outcomes of the choice of this ‘linguistic archetype’ is the somehow fluctuating treatment of word forms that deviates from ‘classical’ Greek. As a tacit rule, what is in fact a ‘linguistic variant’ with respect to classical Greek is assumed to be the ‘regular’ form, in a more or less conscious consideration of the cultural and linguistic environment of the papyri from Graeco-Roman Egypt. Nevertheless, the situation is not that clear, and sometimes we do find sporadic editorial ‘regularizations’ that do not relate to outright scribal mistakes, as traces of

---

* This lecture was first delivered in Trier on June 30, 2016, in the framework of the “Vorträge im Rahmen des Kolloquiums ‘Probleme des griechisch-römischen Ägypten’”. My grateful thanks to Fabian Reiter for the kind invitation. An updated Italian version has been presented at the “Greek Medical Papyri” conference. The paper falls into the project “Online Humanities Scholarship: A Digital Medical Library Based on Ancient Texts” (ERC-AdG-2013-DIGMEDTEXT, Grant Agreement No. 339828, Principal Investigator: Prof. Isabella Andorlini), funded by the European Research Council at the University of Parma.

1 Cf. Youtie, Criticism, 13-15. For the cases of copies and duplicates, see below.

2 A thorough discussion of this topic can be found in Stolk, Encoding.

3 The most typical example is constituted by the verbal voices of ginomai, Koine form of classical gignomai (For the loss/assimilation of gamma before ny cf. Mayser, Grammatik I, 164-6 [Ptolemaic age]; Gignac, Grammar I, 176 [Roman age]; in the Byzantine age gamma comes back), which in the
inconsistencies in the architecture of the linguistic standard.

The situation gets more and more blurred as we gradually enter the quicksand of the linguistic variations that affect the language of the papyri. Even when we look at the most frequently attested cases in the Greek documentary papyri, we do find persistent uncertainties. Let us consider, for example, the most widespread case of the well-known iotacism, phonetic exchange between iota and epsilon-iota, which “indicates the identification of the classical Greek /ei/ diphthong with the simple vowel /i/”. Out of the 25 occurrences in which the full spelling can be recognized with certainty, 17 exhibit the ‘standard’ form ending in –εια, while 8 feature the iotacized ‘variant’ ἑρμηνια. Among the latter, in 5 cases the editors decided to ‘regularize’ the form by indicating the ‘standard’ reading in the apparatus (two cases are weird in that the re-editions of the papyri get rid of the regularization in the apparatus). In three cases the term is printed as it is, without any further critical annotations, in spite of the existence of strict parallels showing the ‘regular’ spelling.

This had a not insignificant outcome in terms of the texts digitized in the Duke Databank / Papyri.info database, in that the editors’ original reading was retained, generating an evident loss of information in the use of the database. I recently experienced the effects of this shortcoming in person: searching for all the occurrences of ἑρμηνεία, at first I got just a partial result. I have therefore emended the digital texts by adding the ‘regularized’ reading in the apparatus of the three papyri, but the extent of the problem is clear, and not yet solved – note the treatment of the supplied word in BGU I 326,i,1: all the printed editions present the iotacized variant, yet the database encoded the ‘regular’ spelling, and my proposal to fix this has been rejected, as an ‘archetype’ has been preferred to the attested variant.

The following table illustrates the situation:

---

papyrus editions are never noted as variants but treated as standard: cf. Depauw/Stolk, Variation, 213-4. Editorial regularizations seem to occur only when the verb is affected also by iotacism, often in compounds: παραγεινομαι l. παραγίγνομαι in BGU XVI 2651,6; γείνεσθαι l. γίγνεσθαι in Chr.M. 172,i,15; καταγειν l. καταγίγνομαι in P.Bodl. I 17,i,9; παραγεινομαι l. παραγίγνομαι in P.Haun. II 22,5; παραγεινομένων l. παραγίγνομένων in P.Stras. VIII 772 passim. Note the double regularization γίγνεσθαι or γενέσθαι for γείνεσθαι in P.Col. X 280,13 (see below).

4 Cf. Gignac, Grammar I, 189; cases: ibid., 189-91. For the Ptolemaic age, with thorough discussion of the phenomenon, cf. Mayser, Grammatik I, 87-94.

5 Cf. Reggiani, Tradurre.
<table>
<thead>
<tr>
<th>Papyrus</th>
<th>Text</th>
<th>Ed.pr.</th>
<th>Following editions</th>
<th>DDB</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGULARIZED</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P.Kell. I 53,11 (IV AD)</td>
<td>ἑρμηνία</td>
<td>ἑρμηνία, ἑρμηνεία</td>
<td>-</td>
<td>Reg.</td>
<td>-</td>
</tr>
<tr>
<td>P.Lund VI 5 (1),3 &amp; (2),2 (AD 187 &amp; 188)</td>
<td>ἑρμηνίας</td>
<td>ἑρμηνίας, ἑρμηνείας</td>
<td>SB VI 9355: ἑρμηνίας</td>
<td>Reg.</td>
<td>SB is a mere reprint of the ed.pr., but without apparatus; fortunately, the DDB followed the ed.pr.; Bingen, Review, 153 quotes straightforwardly ἑρμηνείας.</td>
</tr>
<tr>
<td>P.Monts.Roca IV 71,3 (AD 141-142)</td>
<td>ἑρμηνίας</td>
<td>ἑρμηνίας, ἑρμηνείας</td>
<td>-</td>
<td>Reg.</td>
<td>-</td>
</tr>
<tr>
<td>P.Thead. 13,ii,15 (AD 321)</td>
<td>ἑρμηνία</td>
<td>ἑρμηνείας, ἑρμηνείας</td>
<td>P.Sakaon 34: ἑρμηνίας</td>
<td>Reg.</td>
<td>Fortunately, the DDB followed the ed.pr.!</td>
</tr>
<tr>
<td>NOT REGULARIZED</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BGU I 326,ii,15 (AD 194)</td>
<td>ἑρμηνία</td>
<td>ἑρμηνία</td>
<td>Chl.M. 316; Sel.Pap. I 85; FIRA II 50; Jur.Pap. 25: ἑρμηνία</td>
<td>Emended</td>
<td>Note at i 1 the supplement ἑρμηνείας in DDB (not in the editions): lectio facilior!</td>
</tr>
<tr>
<td>SB X 10288 (2) [Polotsky, Documents] (AD 132)</td>
<td>ἑρμηνίας</td>
<td>ἑρμηνίας</td>
<td>Lewis, Documents; P.Yadin I 27, 15: ἑρμηνείας</td>
<td>Emended</td>
<td>'Regular' spelling in P.Yadin 11, 30 and 16, 33 &amp; 36 (parallel passages)</td>
</tr>
<tr>
<td>P.Fay. 23,12 (AD 131)</td>
<td>ἑρμηνίας</td>
<td>ἑρμηνίας</td>
<td>-</td>
<td>Emended</td>
<td>'Regular' spelling in P.Oxy. XXVII 2472, 3 where the same official title occurs</td>
</tr>
</tbody>
</table>

So far, so good: we have noted how ‘irregular’ word forms can be irregularly dealt with by modern editors, as well as the not negligible outcomes of such circumstances in terms of digitized texts. Which begs the question: is the choice of a linguistic archetype effective for contexts in which linguistic changes occurred? More generally speaking, can the choice of a linguistic archetype be

---

6 This is the case, for example, with later Egypt, when the use of dative in the Greek language of the papyri experiences a general shift to genitive, on which cf. Stolk, Dative. Thus we can have the case of BGU XIII 2332.20 (AD 375), where ὑπάρχω + genitive (μου) is regularized in dative (μοι) according to the classical use (for more cases of similar variations, see Stolk, Variation, 85 ff.), but also the case of SB XVIII 13947.15 (AD 507), where ὑπάρχω + dative (μου) is regularized in genitive (μου) as if the latter was then the correct form (cf. Depauw/Stolk, Variation, 213; see also Stolk, Variations, 93). Note that this correction occurs in the Sammelbuch, not in the first edition (Sijpesteijn, Papyri, 138).
universally effective for a multilingual society? How much role does the frequency of attestation of a form play? Trevor Evans has recently demonstrated, through examples from the Ptolemaic archive of Zenon, the importance of considering terms of comparison among the papyri themselves in order to conceive a more or less correct idea of linguistic ‘standard’, or better, in his own words, of “substandard usage in documents of the same place and time”:

we should be building our understanding of an emerging standard language in non-literary papyri from this internal evidence much more than from the practices of classical literature,

as he convincingly concludes. Furthermore: how much role do personal consciousness and individual preferences or customary habits play? In Sir Kenneth Dover’s words, “[n]o utterance is such that its author cannot care what it sounds like”, why shouldn’t we care it as well? Should we regularize according to our own linguistic taste, or according to the ancient author’s one? Note that the purpose of textual criticism is to establish what an author exactly wrote, and that, by definition, a linguistic variant is any of the different phonetic, morphological, or graphical aspects under which a word can appear in a language, the choice of which can be due to personal reason and preferences, or to archaic, regional, poetic uses.

Let us take into consideration an interesting example of these ‘substandards’. εἰσοπτρὸν is a quite rare word meaning “mirror”, in parallel with the more frequent κάτοπτρὸν. The ‘regular’ or ‘standard’ form, with the preposition εἰς, is not attested as such in the documentary papyri. It always occurs in linguistic variants that do not correspond, as we might expect, to the iotacised form *ἴσοπτρον, which is, surprisingly, unattested even in the literary texts. The first attested variant, ἔσοπτρον, is in fact quite frequently used in literature too, so that it is registered in the main repertories11 and is treated as the ‘standard’ form, without regularizations, in the papyri, where it occurs six times:

---

7 And for a society at all: but this is another question.
8 Evans, *Standard*, 205.
9 Let us consider what C. C. Edgar called Amyntas’ “weakness” for ἄφεσταλκα: the preference accorded by one of the main characters of the Zenon archive for the aspirated perfect form of ἀποστέλλομαι, instead of the classical ἀπέσταλκα (in the note to P.Cair.Zen. I 59047,1; cf. Evans, *Language*). This is certainly not a regular form, not even a correct one, but what to think when an author uses with a certain constancy such an irregular form? Shouldn’t we assume it as standard (or, according to Evans’ terminology, “substandard”), since it was almost systematically (perhaps consciously?) employed by an author? And shouldn’t we reverse the situation, positing the classical form as a variant of the idiosyncratic spelling? By the way, the idiosyncratic form is not regularized in P.Cair.Zen. III 59435 + P.Cair.Zen. IV p. 289,3.
11 See e.g. Daris, *Spoglio I*, s.vv., where the two forms are recorded separately. LSJ s.v. εἴσοπτρον
The second attested variant is ὀσυπτρον. It occurs in the papyri only, and is phonetically explained as a vocalic metathesis of ἔσοπτρον, being the latter either the unattested iotacised form of εἴσοπτρον or the actual pronunciation of ἔσοπτρον.12 The shift between iota and hypsilon is an unsurprising phenomenon,13 and indeed the spelling ὀσυπτρον occurs once (though it would be better to check the reading, since no picture of the papyrus is available online). Another explanation makes it derive directly from ἔσοπτρον through other typical vocalic changes (/e/ > /o/ and /o/ > /u/);14 in this case, ὀσυπτρον would be a phonetic variant. Anyway, all the seven papyrological attestations of ὀσυπτρον (as well as the only instance of ὀσιπτρον) are treated as ‘irregular’ variants, and regularized sometimes in εἴσοπτρον, sometimes in ἔσοπτρον:15

This is apparently weird, in assuming two ‘regular’ forms for the very same variant. Moreover, it does not take into consideration the possibility, claimed by Isabella Bonati, that ὀσυπτρον may actually have been a standard (or sub-
standard!) form, not only a phonetic variant, being attested as many times as (or even more frequently than) the allegedly standard form ἑσοπτρον, and having produced even a Latin loan word, the *osyptrum* of ChLA IV 249r.12. The critical uncertainties are of course mirrored by the digital uncertainties: the impression is that the extant syntax to encode linguistic variants – which reflects the current scholarly position – is not really designed to support complex cases or potential substandards, and this turns into simplifications that do not help the research.16

When we digitize the text of a papyrus, we must take a decision. Digitizing means indeed encoding the text in a machine-readable language that is conventional, logical, precise and standardized.17 Any possible uncertainty may result in potential loss of information and therefore in limitations to the enormous potentialities of the database. For example, the current markup tag used to indicate a linguistic variation like the iotacism is called *regularization*:18

Leiden+: <ερμηνεία|reg|ερμηνία:>

XML (TEI/Epidoc): <choice><reg>ἐρμηνεία</reg><orig>ἐρμηνία</orig></choice>

HTML: Text: ἐρμηνία App.: ἐρμηνεία

It is evident that behind such a syntax lies the traditional idea that any variant (expressed by the XML tag <orig>, which stands for “original reading”) must be brought back to a form that is assumed as regular (the XML tag <reg>, i.e. “regularized term”).19 This is highly affected by the uncertainty and inconsistency

---

16 If I want to collect all the extant attestations of that word in the corpus, I have to know both ‘standard’ versions – ἑσοπτρον and ἑσοπτρον – and then perform the search for both, being still unable to do proximity searches when the word appears in the apparatus. See also the critical observations by Stolk, *Encoding*, passim.

17 As is by now known, the encoding language in use is called Leiden+, and is an advanced version of the Leiden editorial conventions for printed editions. Some conventions are the same (e.g. underdots to indicate uncertain letters), some other features are encoded through particular ‘tags’ (i.e. labels); both can be interpreted by the platform (the Papyrological Editor) and are automatically converted into a standardized markup format called XML and, at the same time, in a human-readable HTML output that looks very close to a traditional print edition (it is important to stress that the XML code contains semantic information which adds meaning to the text, while HTML is purely aimed at displaying the text in a more or less lovely format). Cf. Reggiani, *Digital Papyrology I*, 234 ff.


19 Indeed, the syntax to tag the outright scribal mistakes (|corr|) works the very same way. For both, see the thorough discussion by Stolk, *Encoding*, to which add Reggiani, *Digital Papyrology I*, 236-7, and Id., *Concept*, 24-5. It is true that the HTML output display on Papyri.info – where initially the ‘regularized’ form was shown in the text and the ‘original’ one in the apparatus, as an earlier *Duke Databank* legacy – follows now a deeper awareness of scribal phenomenology in displaying the two forms the other way around (as from late 2011, as announced by Sosin, *Updates*), but the syntax semantics remains the very same.
in defining a ‘standard’ form, which is affected – in turn – by discussion about the
very nature of linguistic variation: for instance, how should we properly encode
the interlaced relations between ἐϊσοπτρὸν/ἔσοπτρον and ὀσὺπτρὸν?

We just considered a couple of cases pertaining to two infrequent words, poorly
attested in the corpus. But what about more frequent words? Are such regularizations
consistent? First of all, a tool to monitor linguistic variation in the documentary papyri
is needed. This tool is now provided by the Trismegistos Database of Text
Irregularities, a databank that catalogues all the attested linguistic variants in the
papyrus corpus.20 It is a first, partial response to an old need, the one already sketched
by Paul Maas that “[t]o reach firm ground […] it would be necessary to prepare a
catalogue of all peculiar errors”21 of a textual tradition, and developed from the digital
viewpoint by Lorenzo Perilli, who provided some guidelines and desiderata for the
construction of a corpus of philological variants to be connected to the TLG
database.22 The switch from the printed medium, intrinsically limited, to the digital
space, which offers potentially endless possibilities of handling the texts, is a
momentous occasion for rethinking the concept itself of textual variation (of any
kind). On the documentary side, the Sematia platform developed at Helsinki to
support the linguistic annotation of the Papyri.info corpus23 proceeded along a parallel
track in envisaging a “variation” layer (still in progress) that will handle linguistic
variants from an innovative viewpoint.

When turning to the literary papyri, and to the technical corpora that are usually
addressed as ‘paraliterary’ or ‘subliterary’, like medical papyri, philological issues
merge with the said linguistic issues, complicating the picture even more.24 A good
example of the situation is P.Aberd. 124 = GMP I 1 (II cent. AD, <http://
litpap.info/dclp/63334>), which preserves, in the first column of the recto, a
portion of the Hippocratic treatise De fracturis, paragraph 37 (III 540, 16 L.), on
the dislocation of the knee. It is, in its editor’s words, a “textually accurate copy
of the Hippocratic treatise”, which “offers a good Greek text which does not share
the banalities of the vulgate tradition”.25 In this fragment we do find variants

21 Maas, Criticism, 14
22 Perilli, Filologia, 36-50.
23 Cf. Vierros/Henriksson, Preprocessing, and Vierros, Annotation; see also below.
24 The project Digital Corpus of Literary Papyrology (DCLP), held mainly at the Heidelberg
Institute of Papyrology, has developed a database similar to the Papyri.info platform but devoted to
literary and paraliterary texts on papyrus (<http://litpap.info>). The endeavour has been as pathbreaking
as arduous, since Papyri.info was designed to host documentary texts only and simply did not offer
enough tools to treat the special features of literary texts. Many questions are still under discussion, and
some pilot projects have contributed with their attempts and suggestions – namely the “Grammatically
Annotated Philodemus” & “Anagnosis” Projects (Würzburg) and the Parma DIGMEDTEXT Project.
On the former (and the DCLP in general) see Ast/Essler, Corpus, as well as Reggiani, Digital
Papyrology I, 250 ff.; on the latter, see the bibliography in Reggiani, Concept, 8 n. 38.
25 Andorlini, De fracturis, 4-5. It is not rarely the case that papyri preserve more genuine text
already attested in the manuscript tradition: at ll. 4-5 of the papyrus, we read τὸ δὲ μέγα, where the codices (and the editions) have τὸ δὲ τοῦ βραχίονος ἵθρον μέγα. The length of the lacuna at the beginning of line 5 excludes the presence of the second article τὸ, in accordance with the only testimony of manuscript M (Parisinus 2247), which omits it.26 We do find, on the other hand, passages completely divergent from the codices. This is the case with ll. 11-12, where the length of the gap and the shape of the following traces (κῶλμα ἐόν, ἀφ’ οὗ (or ἀφ’ οὗ περ) πέφυκεν, which is of course printed in all the editions, in favour of a previously unattested κόλμα ἐόν, ἀφ’ ὁκ]οιου (or ἐκείνου) πέφυκεν.27 The fragmentary state of the papyrus makes things slightly more difficult, but it is clear enough that we are not dealing with the known tradition of the text. A last example is even more interesting, as it merges philological and linguistic issues in one single word: at line 14, where the codices (and the editions) have the ‘regular’ Ionic form πήχεος, the papyrus shows clearly (as already transcribed by Eric Turner, ed.pr.) πήχεως, the Koine form,28 which is an ‘interference’ (generating a variation) of a typical form pertaining to the language of the documentary papyri (where, on the contrary, it would be the standard form, as we saw earlier).

How to ‘regularize’ such occurrences? Is it even possible to speak of ‘regularization’ at all? In fact, not rarely does linguistic variation, in technical texts especially (but not only!), bear broader cultural significance.29 On the phonetic level, ‘substandard’ forms – equivalents to the abovementioned cases of ἕρμης(ε)ία and ἐ(ι)σοπτρον/ὄσποτρον – very often betray cultural interferences that deserve more care than distinguishing between ‘regular’ and ‘irregular’ forms. Even a ‘simple’ case such as the fluctuation between σμύρνη and ζμύρνη “myrrh” can generate hesitations: according to the grammar, σμύρνη should be the ‘regular’ spelling, while “σ is frequently replaced by ζ before a voiced consonant”30 – in the medical papyri indeed, in particular, the term “è scritto quasi regolarmente ζμ.”.31 It is really very difficult to choose a ‘regular’ form, to which trace a ‘variant’ spelling back: for example, though Gignac notes that “σμ- spellings are found especially in Byzantine papyri”,32 plenty of late antique medical papyri the word is in fact spelled out with

---

26 Andorlini, De fracturis, 6 and note ad loc.
27 Andorlini, De fracturis, 5 (“The possibility of a textual variant must be reckoned with”) and note ad loc. for detailed discussion.
28 Andorlini, De fracturis, 6 and note ad loc. with further bibliography on the inconsistency in the Ionic dialect employed by the Hippocratic writers and the papyrus tradition.
29 Another interesting aspect of these textual fluctuations is that of hypercorrections, which will be dealt with (as regards medical papyri) in Maravela/Reggiani, Scribes.
30 Gignac, Grammar I, 120 (121-2 on σμόρνη). For the Ptolemaic papyri, see Mayser, Grammatik I, 204.
31 Andorlini, Ricette, 61 n. 54.
32 Gignac, Grammar I, 122.
ζμ to the extent that it turns into no less than a monogrammatic symbol featuring the character zeta in plain view (צ, in P.Acad. inv. 4.25 [V AD]), configuring itself as a proper ‘substandard’ spelling.

There are more complex instances. P.Oslo inv. 1576, a fragment of a catechism dealing with tumour-like diseases, partly overlaps with the text of P.Oxy. LXXX 5239 (both II-III cent. AD). The scarceness of the surviving portions of text makes it hard to say whether the questionnaire derives from the treatise or they are two different outcomes of a same ascendant. As far as the extant parallel text is concerned, the wordings diverge from each other only for one variant: ὑδροκήλη (P.Oslo, l. 5) vs ὑγροκήλη (P.Oxy., l. 15). The latter is usually considered as a minority variant (LSJ, quoting Poll. IV 203) of the former, used e.g. by Ps.Gal. Def.med. 424 = XIX 447,12-13 K., but it is in fact attested three times among the medical writers. Are we facing a trivialization in the Oslo papyrus, or a simple phonetic variant in the Oxyrhynchus papyrus, or just two different traditions bearing the same degree of ‘correctness’, attesting to a fluid notion of technical language? Moreover, in the following line of the Oslo papyrus (not paralleled by its Oxyrhynchus counterpart any more) we read ἐρυτρ[οειδῆ, which looks like a phonetic variant of ἐλυτροειδῆ “lid-like”, “cover-like”, attribute of one of the membranes enveloping the scrotum. Rho for lambda is indeed a very frequent phonetic exchange in the language of the Greek papyri, but the same variation is to be found among the manuscripts preserving Ps.Galen’s Introductio seu Medicus, containing a descriptive passage (XIV 719,5-10 K.) of the same anatomical part, making it quite hard to establish degrees of ‘regularity’ in the spelling employed.

On the morpho-syntactic level, we witness phenomena that go far beyond case variation as thoroughly examined by Joanne Stolk. In medical prescriptions and

33 See P.Michael. 36 (Byz. age); GMP I 14 = P.Sijp. 6 (IV/V AD); GMP II 8 (V AD); SB XXVIII 17138-17139 (V); MPER n.s. XIII 8 (second half V AD); SB XIV 11964 (V/VI AD); P.Cair.Masp. II 67141 (VI AD).
34 On this monogram cf. the observations by Fournet, Papyrus, 319-20. The papyrus is republished in Fournet, Bibliothèque, 185-7.
35 Maravela/Leith, Catechism. The papyrus will be republished in the forthcoming P.Oslo IV. I am most grateful to Anastasia Maravela for sharing her drafts of the new edition and for discussing with me some textual and linguistic details.
37 Cf. Gignac, Grammar I, 105.
38 The previous editors corrected it in ἔφρυτροποῦς, but the newest Belles Lettres edition (Petit, Médecin) prints ἑλυτροειδῆς (XII 11, p. 40,1; see Petit, Médecin, XCVI-XCIX for the description of the manuscript tradition). Quite interestingly, the author of the treatise came possibly from Egypt (cf. Petit, Médecin, 1-L), which suggests that the phonetic variation could have worked both ways. I discuss this and the preceding case also in Reggiani, Concept, 28, and Id., Literacies.
39 For other similar cases cf. Reggiani, Concept, 26-7.
40 Stolk, Variations; Dative; Encoding. See the case of ἴπαρχος mentioned above.
recipe-like texts, it is remarkably frequent the use of the verb χράομαι in the
imperative form χρῷ “use” to introduce specific instructions about the
composition of medicaments (typically, “the final formula of a recipe which
suggests how an ointment should be mixed and applied”). This is typically
accompanied by the indication of the substance to which the previous compound
is to be mixed (e.g. SB VIII 9860,ii,9: χρῶ ἐν ὠάδι “use in/with water”; P.Tebt.
II 273 = GMP II 5,ii,13: μὲς ὠίνου χρῶι “use with wine”) or of the ingredient
to be used (e.g. P.Oxy. VIII 1088,i,19: ὑαλκίτηι λήᾳ χρῶι “use pounded rock-
alum”; P.Oxy. LXXIV 4975 (1),4: τῇ σποδῷ χρῶ “use the powder”).
Nevertheless, it is not rarely the case that the syntagm “use with water” appears
under the a-syntactic aspect ὕδωρ χρῷ (e.g. P.Tebt. II 273 = GMP II 5,iv,5, vii,17,
viii,5, 22; P.Princ. III 155v,9), which goes far beyond the apparent ‘incorrect’
anacoluthon, becoming a distinctive mark of medical recipes. It would be rather
senseless to ‘regularize’ such peculiar circumstances, for which we may well
speak of ‘formulaic substandards’, which increasingly tend to detach from the
syntactic architecture of the discourse and to constitute textual and graphical units,
completely released from the context. An interesting case – not as well-
established as ‘ὕδωρ χρῷ’ and therefore perhaps even more significant – is also
given by P.Princ. III 155r,7, where the common instruction “use with wine” is
written as οἴνου χρ (ῶ), in which the genitive is interpreted by Isabella Andorlini
as influenced by the genitive endings of all the ingredients listed above (ll. 2-6).
Another possibility is that the syntactic flaw is a formulaic derivation of the
‘regular’ με‘ οἴνου. The absence of parallels prevents us from speaking of
‘formulaic substandard’, but the overall cultural context encourages us to be
particularly cautious when dealing with otherwise apparent ‘irregularities’.
From the digital viewpoint, the embarrassment in handling such linguistic fluctu-
ations, in which it is not actually possible to identify a ‘correct’ or even a ‘standard’
form as opposed to a ‘deviant’ hypostasis, adds to the difficulties raised by the
treatment of philological cases. In the Aberdeen example presented above, a first

---

42 The Princeton papyrus is republished by Andorlini, Prescription, 6-11 in this volume.
43 Cf. Andorlini, Gergo, 163: “I due vocaboli […] non mantengono alcun legame sintattico con
quel che precede, né tra di loro”; also Andorlini, Prescription.
44 Cf. Andorlini, Gergo, 160-1 n. 39, apropos of PSI XXI Congr. 3: “[n]ella formula […] l’impe-
rativo pres. di 2a persona del verbo χράομαι (χρῶι), scritto per esteso e con iota ascritto, e non
abbreviato nel consueto monogramma dei testi di età romana (⫷) […] , è avvertito ormai come
asintattico, e si avvia ad assumere la funzione di sigla di chiusura, svincolata dal contesto”; cf. ibid.,
163. On the monogram cf. Gazza, Prescrizioni II, 111. On the reduction of the entire formulaic
syntagm to a couple of symbols – which stresses its autonomy from the context – cf. Andorlini,
Gergo, 163. Sometimes, it is the only word ὕδωρ on its own that is used as a formulaic marker at
the end of the prescription (see e.g. SB XXVIII 17139,27).
45 Andorlini, Prescription, 8.
attempt to cope with the latter has been to use the syntax for the editorial correction (tag |ed|), which is intended to encode the modern interventions on previous editions; the result was nice from the viewpoint of the display, but the semantics were completely incorrect. Another way in which we could encode variants is the syntax for the editorial alternative readings (tag |alt|), designed to indicate two or more possible readings for the same text. This might be semantically closer to our intention, but the search functions would be limited; moreover, if I wanted to add information on the author of an alternative reading, I cannot use this tag, because this function is not supported (as it is, conversely, by the |ed| tag). Currently, no tool is available to deal with such cases in a proper way. The possibility of adding a line-by-line commentary can be exploited for this purpose (as we did), but then the information will be searchable with many limitations.

Let us move to an even more puzzling problem. P.Tebt. II 272, verso (late II cent. AD, <http://litpap.info/dclp/60048>), is a fragment of Herodotus Medicus’ De remediis, describing the symptomatology of thirst and its treatment; the text corresponds in part to an excerpt of Herodotus Medicus preserved with Oribasius’ treatment of thirst in case of fever (Coll.med. V 30,6-7 Raeder = CMG VI 1,1). At l. 5, where the text reads αἰτίαι τῆς προσφορᾶς introducing the different reasons for giving the sick something to drink, the scribe adds two groups of three letters between dots above the line: *τῶν* above τῆς, and *ρῶν* above ρᾶς.

46 For both editorial tags cf. Reggiani, Concept, 43-4.

47 For example, let us consider a case of iotacism that conceals a double alternative reading: in P.Col. X 280,13 γείνεσθαι can be ‘regularized’ as either γίγνεσθαι or γενέσθαι. This is encoded as follows: Leiden+: <:γίγνεσθαι|altern>γείνεσθα̣ι>: XML: <choice><reg><app type="alternative"><lem>γίγνεσθαι</lem><rdg>γενέσθαι</rdg></app></reg><orig>γείνεσθα̣ι</orig></choice>. HTML output: ταῦτα γείνεσθα̣ι (text); app: l. γίγνεσθαι (or γενέσθαι). This allows one to perform a proximity search only for the word printed in the text, i.e. the ‘irregular’ one, so that if I search for passages in which ταῦτα is followed by a word containing γειν- within 3 characters of distance, I will find our papyrus (along with another one). Instead, if I search for ταῦτα followed by either γιγν- or γεν- I will not find our papyrus any more.

48 Documentary papyrologists seem to have some kind of allergy to philological issues. When we have to do with duplicate documents – which would allow them to experience the thrill of a textual collation, and therefore to develop strategies to cope with textual variants – it is usually published one copy (best if the better preserved one), stated that another copy of the same text does exist, and that is about it. Consequently, no tool has been developed to treat digitally such circumstance: an earlier tag, introduced in the Duke Databank of Documentary Papyri to indicate an alternative reading from a duplicate of the same text (cf. Willis, Data Bank, 169-70), has been completely abandoned. On copies and duplicates, which fall into the framework of genetic criticism, see Cribiore, Criticism (in this volume); Jördens, Entwurf; Mirizio, Antigrapha; Ead., Archetypes; Nielsen, Catalog; Stoop, Copies, 185; Yuen-Collingridge/Cheat, Copyist; for their digital treatment, cf. Reggiani, Digital Papyrology I, 268 n. 57, and Id., Concept, 35 n. 144. On the philological side, the proposal to use a tag |var| raised some practical and theoretical issues, on which cf. Reggiani, Concept, 35 n. 145.

49 I thank very much Todd M. Hickey and Derin McLeod for the help in getting a high-resolution picture of the fragment.
When digitizing that, we must find a way to state that:

1) the scribe juxtaposed some text above the line, but not in the form of an addition \textit{supra lineam} or \textit{infra lineam}, since it is clearly a variant of the syntagm below (plural instead of singular). We cannot therefore use the standard way to tag supralinear or interlinear insertions, since it would not make sense.

2) He actually wrote *τῶν* and *ρῶν*, but meaning \textit{τῶν προσφορῶν}: we must report both readings, since we have to represent what exists on the papyrus but also to be able to search for the full combination of words, and hopefully to find it. For both of these reasons, therefore, we cannot just encode τῶν and ρῶν by themselves.

3) Since nothing appears deleted, it is not clear if the scribe wanted to correct the text or just juxtapose two different versions of the same passage. In fact there are examples of philological corrections added \textit{supra lineam} without deletion marks, as well as of deletions indicated by means of overdots (both features are attested in P.Oxy. XXIV 2404, ii 6).\textsuperscript{50} Otherwise, writing a word between dots can be a way to highlight a correction added later on (see e.g. the \textit{koppa} in P.Eirene III 25,3, with comm. \textit{ad loc.}).\textsuperscript{51} We cannot be sure of what is going on here because this variant is unattested in the manuscript tradition, i.e. in Oribasius’ passages quoting Herodotus Medicus, which all have the singular form. We would have a scribe correcting the form unanimously preserved by the manuscript tradition and replacing it with an unattested variant. The P.Tebt. editors speak of “correction or alternative reading,” Marie-Hélène Marganne of “hésitation.”\textsuperscript{52} If we should define it, we ought to call it a ‘scribal variant’. And once we define it like that, one main question arises: which is the ‘archetype’, and which the ‘variant’? Once more, we may describe the situation in the line-by-line commentary, but the text itself will lose useful information. The current \textit{Papyrological Editor} is clearly not trained to treat such cases in a clear and effective way.

The solution I proposed is no more than a trick that exploits the tags intended for editorial corrections and regularizations, with some descriptive comment added:

\begin{verbatim}
LEIDEN+

4. [τοῖς τῆς ἐπιδόσεως χρόνοις <πλεί 
5.- [ους̣ αἰτίαι |ed| πλεί5.- [σ]ται <αι>τίαι > <τῆς προσφορᾶς=actual text|ed|*dot*τῶν*dot*<:προσφορῶν|reg|*dot*ρῶν*dot*:>=scribal original insertion above line: > εἰσιν·

\end{verbatim}

\textsuperscript{50} Cf. Colomo, \textit{Osservazioni}, 24.
\textsuperscript{51} On dots as lectional sigla cf. McNamee, \textit{Sigla}, 23 (not relevant to us).
\textsuperscript{52} Marganne, \textit{Fragment}, 76.
XML

<milestone rend="paragraphos" unit="undefined"/>
<lb n="4"><supplied reason="lost">τοῖς τῆς ἐπιδόσεως χρόνοις <app type="editorial"><lem>πλεί</app></supplied></lb><lb n="5" break="no"><supplied reason="lost">ou</supplied></lb><supplied reason="lost">τῆς ἐπιδόσεως χρόνοις πλείς αἰτίαι <app type="editorial"><lem>πλεῖσ</lem><rdg>πλεῖσται αἰτίαι</rdg></app> τῆς προσφορᾶς ἐσιν·</supplied>

HTML

[toῖς τῆς ἐπιδόσεως χρόνοις πλείς αἰτίαι τῆς προσφορᾶς ἐσιν·
5 οὐς αἰτίαι(*) τῆς προσφορᾶς(*) εἰσιν·

App.:
4-5. πλεῖσται αἰτίαι<sup>prev. ed.</sup> τίς προσφοράς<sup>scribal original insertion above line</sup>

It is clearly a tightrope walking, semantically inconsistent, but Leiden+ and TEI/Epidoc XML do not offer anything better for the moment.

Some test cases provided by Federico Boschetti,54 in the attempt to provide the literary databases such as the TLG with the due critical apparatuses, might direct our steps towards the enhancement of the Leiden+ syntax with tags specifically oriented to treat textual fluctuations, either philological or linguistic. Basically, this would mean to connect more information to the text encoded in the database.

What follows is a simple specimen from Boschetti’s test case:

197 ἡ 819 δ᾿ 820 ἔσφαδζε, καὶ 822 χεροῖν 823 ἐντη 824 δίφρου 825
197. αὐτὴ δίφρον Canter.

<itm>
<vrs>197.</vrs>
<rdng><g pos="824">αὐτὴ</g> <g pos="825">δίφρον</g></rdng>
<resp>Canter</resp>.
</itm>

53 The text has been digitized by student Clara Quartarone during a workshop at the University of Parma.
54 Boschetti, Methods, 4 ff.
Each word is identified by a unique number; the variants are aligned (i.e. linked) word by word; further information such as the origin of the alternative reading is provided within its own tag (this is not without issues, such as the problem of how to tag broken words).

A significant improvement in the addition of meaningful information comes indeed from linguistic annotation, a powerful methodology developed by corpus linguistics, the branch of linguistic studies that deals with corpora of texts as representative samples of an entire language. Annotating a corpus means to tag textual elements in a systematic way, adding some kind of linguistic information.\(^{55}\)

It allows to describe, record, interpret and analyse linguistic information at several levels, in which each layer corresponds to a particular category of relevant information that is made available for quantitative analyses and deep search options by means of regular expressions (i.e. combinations of search parameters) and/or XML query strategies. For example, a part-of-speech annotation layer (the one usually known as treebank) connects each token (basic unit of the linguistic corpus) with information about its morpho-syntactic aspect, allowing investigations on the lexical, phraseological, syntactic pattern of the reference language.\(^{56}\) Among the others, an annotation layer devoted to variant tagging would significantly improve our textual database.\(^{57}\)

An annotation layer devoted to textual fluctuations may find a broader application within the concept of transtextuality, as it has been recently analysed by Monica Berti with reference to historiographical fragments.\(^{58}\) Transtextuality defines the various possible relations among texts,\(^{59}\) and we may refer it not only to a network of quotations and parallel passages (“intertextuality”),\(^{60}\) but also to the aspect of fragment which very often the papyrus, be it literary or documentary, shows.\(^{61}\) Indeed, we can define the fragmentary character of the papyrus as a sort of ‘non-voluntary quotation’, selected by the chance and by the material circumstances rather than by an author’s will. The transtextual link will work, in this case –

\(^{55}\) To some extent, the Leiden+/XML markup is a kind of non-linguistic, rather semantic annotation.

\(^{56}\) For an overview of linguistic annotation applied to papyrus texts see Reggiani, Digital Papyrology I, 178-85. For specific projects see Riaño Rufilanchas, Philodemus, and Essler/Riaño Rufilanchas, AristachusX (Herculaneum papyri); Vierros/Henriksson, Preprocessing, and Vierros, Annotation (documentary papyri); Porter/O’Donnell, Phenomena (papyrus letters). I dealt with the topic of linguistic annotation applied to the corpus of the Greek medical papyri in Reggiani, Project (pp. 344 and 347 on variant/error tagging), Perspectives, 2-3, and eventually Concept, 18-27.

\(^{57}\) Stolk, Encoding, 133-6, persuasively suggests that annotation of linguistic variations should pass through lemmatization of the terms involved.

\(^{58}\) Cf. Berti/Romanello/Babeu/Crane, Collecting; Berti, Citazioni, 442-7; Ead., Texts; Ead., Representing.

\(^{59}\) Cf. Reggiani, Concept, 28, with bibliography (namely, with reference to Gérard Genette’s textual theory).

\(^{60}\) Cf. Reggiani, Concept, 33 ff., with bibliography.

\(^{61}\) Cf. Reggiani, Concept, 16-17.
following the specific terminology –, as a relationship of ‘hypertextuality’, which
describes the derivative connection of a ‘hypertext’ (in our case, the original
document, lost, more or less recoverable in a philological way) with a ‘hypotext’
(our fragment), showing various degrees of interference: physical and mechanical
ones, but also scribal mistakes, linguistic fluctuations, and philological variants.62

As with the digital encoding of fragmentary quotations,

hypertextual models allow to rethink the fundamental question of the relation
between the fragment and its context, representing and expressing every
element of print conventions in a more dynamic and interconnected way.63

This perspective would bring innovative solutions to the current question of how
to manage the annotation of variants in the papyri. If we cease to consider a variant
as a ‘deviant version’ to be ‘corrected’ and ‘regularized’, we can overcome the
deadlock by looking at the full set of variants as a network or a system, and by
thinking the digital edition as a multitext, a place for a dynamic collation of several
editions (or, in our case, versions), stratified in the time (or even at the same time
level). As Monica Berti puts it,

collecting multiple critical editions of the same text means building a
‘multitext’, which is a ‘network of versions with a single, reconstructed root’,
so that scholars can compare different textual choices and conjectures produced
by philologists. This process involves a new way of conceiving literary
criticism because it produces a representation and visualization of textual
transmission completely different from print conventions, where the text that is
reconstructed by the editor is separated from the critical apparatus that is printed
at the bottom of the page. In addition, the inclusion of images of manuscripts,
papyri, and other source materials allows the reader to have a dynamic
visualization of the textual tradition and to perceive the different channels of
both the transmission and philological production of the text that is usually
hidden in the static, concise, and necessarily selective critical apparatuses of
standard printed editions. Producing a multitext, therefore, means producing
multiple versions of the same text, which are the representation of the different
steps of its transmission and reconstruction, from manuscript variants to
philological conjectures. This process has fundamental consequences for the
study of ancient sources in general and for fragmentary ones in particular, given
that, while studying fragments and evaluating their distance from the original
version, it is imperative to examine the manuscript variants of the source text,

62 Cf. Reggiani, Concept, 33 ff. (hypertextuality) and p. 45 n. 187 (fragments as ‘virtual’ hypotexts).
63 Berti, Texts, 1.
in order to see what can be attributed to the witness or to the transmission of the text across centuries.\textsuperscript{64}

It is, as has been already pointed out elsewhere,\textsuperscript{65} a completely new and different way of considering critical editions of ancient texts. The texts – in our case, the papyri – becomes meta-texts (meta-papyri) and the critical apparatus dissolves in a network of references, connections, and versions.\textsuperscript{66} This is extremely valuable under the philological circumstances of uneasy textual relations (the ‘living’ medical texts are a perfect example of this)\textsuperscript{67} as well as in the linguistic cases of fluctuations that cannot actually be handled in terms of plain ‘regularizations’. This kind of textual network has also much to do with the ancient philological care as testified by numerous instances (P.Tebt. II 272 verso, decribed above, is just one example): as I contended elsewhere, ancient ‘philology’ – at least in technical fields such as medicine – was much more interested in a fluid textual transmission (the “accretive model of composition”\textsuperscript{68} claimed by Ann Hanson to define the writing history of certain medical texts) rather than in the fixation of a stable (and static), canonical (and constrained), ‘correct’ version of the texts.\textsuperscript{69}

The pathway is still long and perilous, but the first moves have been made: I noticed with pleasure several papers dealing with the treatment of linguistic variants, or related topics, delivered at the latest International Congress of Papyrology at Barcelona (August 2016). I think that from whichever viewpoint one affords the question, the following statement by Greg Crane must unavoidably be kept in mind:

[i]n a digital age, philologists need to treat our editions as components of larger, well-defined corpora rather than as the raw material for printed page layouts.\textsuperscript{70}

\textbf{Bibliography}


\textsuperscript{64} Berti, \textit{Texts}, 4-5.
\textsuperscript{66} The best way of depicting this framework is the ontology model: cf. Reggiani, \textit{Concept}, 1-4 and 45-52.
\textsuperscript{67} Cf. Reggiani, \textit{Concept}, esp. 33-43; also Id., \textit{Literacies}.
\textsuperscript{68} Cf. Hanson, \textit{Fragmentation}, e.g. 305.
\textsuperscript{69} Reggiani \textit{Transmission}.
\textsuperscript{70} Crane, \textit{Editors}.


Colomo, Osservazioni = Daniela Colomo, Osservazioni sullo scriba ossirinchita dell’omega quadrangolare (Johnson A2), “Segno & Testo” 6 (2008), 3-34.


