

A short note on typesetting Latin verse scansion with \LaTeX and Lua \LaTeX

Antoine Bossard

Abstract

Large parts of the Latin literature are written in verse: Virgil's *Æneid* and Ovid's *Metamorphoses* are two well-known examples. As can be noticed from a glance at most specialised textbooks, typesetting the scansion of Latin verses is not trivial. To this end, analysing the syllable quantity, or duration, is required – this is prosody. Conventionally, prosody is denoted with diacritical marks, typically above the vowel of a syllable. Although \LaTeX provides such a feature, it is not always simple to adjust it to authors' needs. In this short note, several issues arising from the scansion of Latin verses in both \LaTeX and Lua \LaTeX are discussed and technical solutions are described.

1 Introduction

Latin verses are based on prosody, that is the syllable quantity, or duration (long, short or common), which in turn defines metrical feet: for example, the foot called dactyl consists in one long syllable followed by two short. The quantity of a syllable is determined by the quantity of its vowel (or vowels, typically in the case of diphthongs).

Scansion of a Latin verse is about identifying feet, which requires checking the quantity of all the syllables. To this end, diacritical marks are used: conventionally, a breve (˘) denotes a short vowel, a macron (¯) a long one and both diacritics used atop one same vowel indicate that the quantity of the corresponding syllable is common.

The verse is usually split into two with a *cæsura*, which is also typically indicated at scansion. While feet are conventionally separated by vertical bars (|), the verse *cæsura* is often materialized with a double vertical bar (||). Elision is sometimes denoted with square brackets or parentheses, and sometimes with a lengthened breve below the space that follows the elided syllable. On the opposite, a hiatus is sometimes denoted with a lengthened inverted breve below the space that separates the two words forming the hiatus [7].

From various Latin grammar textbooks and dictionaries, it can be deduced that scansion typesetting is a non-trivial issue. For example, diacritical marks typesetting is uneven and sometimes lacking in classic textbooks such as [3, 5]. The 2018 edition of [6] is amongst the nicest. More recent textbooks often keep scansion and prosody information at a

minimum [1, 2, 4] and their typesetting sometimes remains below par [8].

In this short note, compiled with \LaTeX and not Lua \LaTeX , we review how to typographically realise the scansion of a Latin verse in \LaTeX and Lua \LaTeX , and the possible shortcomings of these two solutions. A few technical details are given in appendix.

2 The case of \LaTeX

With respect to single diacritical marks applied to single letters, no problem whatsoever has been witnessed when relying on \LaTeX . In both lower and upper case, both accented Unicode characters (e.g. U+0103 ä, U+0100 Ā) and letters manually marked with the commands $\backslash u$ (breve), $\backslash =$ (macron) (e.g. $\backslash u\{a\}$, $\backslash =A$) are correctly rendered. In the case of manual marking applied to the vowel *i*, the dotless version of the letter is of course highly desirable: for instance, $\backslash =\i$ instead of $\backslash =i$ to obtain \bar{i} instead of \dot{i} .

Regarding the stacking of several diacritical marks atop one letter, typically to denote a syllable whose quantity is common, it remains difficult with \LaTeX : $\backslash u\{\backslash =a\}$ is of no avail as it produces \tilde{a} . Relying on additional packages is probably the best solution, like *tipa* by R. Fukui, which notably provides the $\backslash u=$ command to combine a breve above a macron, *stackengine* by S. B. Segletes or *covington* originally by M. A. Covington with its $\backslash twodias$ command.

One single syllable sometimes include several, consecutive vowels: diphthongs, like *æ*, *au*, *eu* and *œ*. The quantity of such syllables is always long. So, when denoting prosody, a macron is expected on top of these vowel combinations.

The rendering of *æ* and *œ* with either manually set diacritical marks or the corresponding accented Unicode characters is acceptable: $\bar{æ}$, $\bar{Æ}$, $\bar{œ}$, $\bar{Œ}$, albeit with the macron being somehow too short. The duration of the *æ* diphthong can however be short in some words, like *præustus*: $\breve{æ}$, $\breve{Æ}$ are rendered as expected. It is however more difficult to add a diacritical mark above the other diphthongs, such as *au* and *eu*, and in the case of *synæresis*, like in *deest* and at the first syllable of *deinde* (a disyllabic word). Relying on the $\backslash overline$ command is one solution: $\bar{a}\bar{u}$, $\bar{e}\bar{e}$. The bar is this time slightly too long and is, more or less depending on the font, too close to the letters. Whilst the former issue can be addressed with $\backslash mkern$: $\bar{a}\bar{u}$, $\bar{e}\bar{e}$ (helpful, by the way, for a lengthened macron too: $\bar{æ}$, $\bar{Æ}$, $\bar{œ}$, $\bar{Œ}$), the latter remains.

Finally, it also happens that two consecutive vowels count as one single short. This is the case, for example, of *ua* in *genua*, a disyllabic word. To avoid

complications, publishers often resort to typesetting *genuă*, which can be misleading during scansion. It could be typeset *genvă* for facilitated scansion, but it is not satisfactory either (the letter *v* notably tends to be avoided in favour of the letter *u* in modern Latin text editions). The diacritical mark could also be moved so that it stands between both vowels: *genuă*, but once again we find this solution not optimal (the mark is too narrow). The `\textasciibreve` command combined with `\llap` is not really helpful either: *genuă*. The Comprehensive L^AT_EX Symbol List by S. Pakin does not mention any extensible breve; note that the `stix` package has a `\widecheck` though.

This discussion on a breve over two letters also applies to the stretched breve sometimes used to denote elision: although a `\textasciibreve` lowered below the base line (and negatively kerned) produces acceptable results, like `ego_ipse`, there is still room for improvement. Similarly, the inverted stretched breve to denote a hiatus can be obtained with a lowered and negatively kerned `\newtie`: `modo_ipse`. Acceptable but not quite satisfactory.

3 The case of LuaL^AT_EX

Support for diacritical marks, both single and multiple, strongly depends on the selected font, and this can rapidly become problematic: the desired font may not support them. In addition, a font that works flawlessly with L^AT_EX is not guaranteed to work with LuaL^AT_EX; this is the case, for instance, of T_EX Gyre Pagella (loaded with the package `tgpagella`).

This is especially the case when adding diacritical marks on capital letters. Besides, this can happen whether Unicode includes the desired character, like U+0232 *Ÿ*, or not, like *Ÿ* (both letters are incorrectly rendered; they have been rendered separately with LuaL^AT_EX and its default font).

Assuming a font that is aware of such diacritical issues has been selected, LuaL^AT_EX brings significant advantages over L^AT_EX as far as we are concerned (cf. Section 2). First, it enables multiple diacritical marks on a character: *ã*, *ÿ* (both letters are correctly rendered; they have been rendered separately with LuaL^AT_EX and the Noto Serif font).

Second, the macron over diphthongs can be further improved: not only its width but also the vertical gap between it and the corresponding letters can be adjusted thanks to the `\Umathoverbarvgap` command. The previous *āu*, *ēē* respectively become *āū*, *ēē* (note the increased gap below the macron; both digrams have been rendered separately with LuaL^AT_EX and its default font).

Next, we acknowledge that Unicode does provide glyphs that could be helpful for our purpose and review their current support with the multiple diacritical mark-aware Noto Serif font. (Note that the results could differ with another font. The support of Unicode by Noto Serif is complete enough so that it is a good candidate for this experiment though.) Once again, the following examples have been rendered separately with LuaL^AT_EX.

First, there is the Combining overline glyph (U+0305), whose purpose is to render a continuous line above several letters. This could apply in our case to a macron added to diphthongs. So, we have applied a Combining overline to each vowel of the *au* diphthong, but the result is not satisfactory: *āū*. In fact, the two overlines are not combined at all.

Second, there is the Combining double macron glyph (U+035E), whose purpose is obvious from its name, and which could once again apply to the macron added to a diphthong scenario. So, we have inserted a Combining double macron between the two letters of the *au* diphthong, but the result is still unsatisfactory: *āū*. Although spanning both letters, the macron is oddly positioned (not centred), and too short.

Third, there exists the Combining double breve glyph (U+035D), whose purpose is to add a breve above two letters. This could apply to two consecutive vowels counting as one single short, as in *genua*. So, we have inserted a Combining double breve between the two letters of the *ua* synæresis, and this time we find the result satisfactory: *ūā*.

Next, it is now possible to rely on the Undertie glyph (U+203F) to typeset a stretched breve to denote elision, which is just fine: `ego_ipse`. Note that in this experiment with the Noto Serif font, the Undertie glyph was unexpectedly “combining” (i.e. like the Combining breve glyph U+0306) and not “spacing” (i.e. like the Breve glyph U+02D8), that is behaved like the Combining double breve below glyph (U+035C), which seems to contravene the Unicode standard. We thus combined the Undertie glyph with two spaces as a workaround.

Finally, there exists the Inverted undertie glyph (U+2054) which is just fine to denote a hiatus between two words: `modo_ipse`. And just like with the Undertie glyph, in this experiment with the Noto Serif font, the Inverted undertie glyph was unexpectedly “combining” instead of “spacing”; we thus combined it with two spaces to fit our needs.

4 Summary

Regarding L^AT_EX, the support of single diacritical marks applied to a single letter is fully satisfactory.

With respect to “native” diphthongs such as æ and œ, macrons and breves are rendered as expected, although the macron may seem a bit too short. Regarding other diphthongs, like au and eu, and other digrams induced by synæresis, rather simple commands could do the trick but results can remain imperfect. In fact, in our experiments, while the macron width is all right, it is placed too close to letters and this is not easily adjustable: a completely new user definition is required if such an adjustment is needed. The breve diacritical mark is not extensible, so the result is only average. The most problematic issue arises from double diacritical marks: their support requires an additional package.

Regarding Lua \LaTeX , Unicode provides useful features for typesetting Latin verse scansion. However, their support strongly depends on fonts, which can provide full, partial (e.g. Noto Serif) or no support at all (e.g. the default Lua \LaTeX font). For instance, the default Lua \LaTeX font does not render correctly multiple diacritics nor diacritics on capital letters. For that matter, the Noto Serif font is significantly better.

Finally and on a side note, depending on the font, the `\textbar` command may produce a vertical bar with too much space around, especially when used within a word: negative kerning will do in such a case.

5 Epilogue

We conclude this note by giving sample Latin verses with their scansion:

ille pedum melior motu fretusque iuventa,
illē pē|dūm mē|li|ōr || mō|tū frē|tū|sque jū|vēntā,

hic membris et mole valens; sed tarda trementi
hīc mēm|brīs ēt | mō|lē vā|lēns; || sēd | tārdā trē|mēntī

genua labant, vastos quatit aeger anhelitus artus.
gēnūā lā|bānt, vās|tōs || quātīt | āegēr ān|hē|lītūs | ārtūs.

multa viri nequiquam inter se vulnera iactant,
mūltā vī|rī nē|quīquam||īn|tēr sē | vūlnērā | jāctānt,

(Virgil, *Aeneid*, book v, lines 430–433)

References

- [1] P. Boehrer, M.F. Delmas-Massouline, et al. *Bled latin*. Hachette, Paris, 2018. ISBN 978-2-01-170040-7. In French.
- [2] B. Bortolussi. *Bescherelle Grammaire du latin*. Hatier, Paris, 2021. ISBN 978-2-218-93175-8. In French.
- [3] A. Cart, P. Grimal, et al. *Grammaire latine*. Nathan, Paris, 2014. ISBN 978-2-09-171242-0. In French.
- [4] F. Gaffiot. *Le Grand Gaffiot Dictionnaire Latin–Français*. Hachette, Paris, 3rd ed., 2014. ISBN 978-2-01-166765-6. In French.
- [5] R. Morisset, J. Gason, et al. *Précis de grammaire des lettres latines*. Magnard, Paris, 1963. ISBN 978-2-210-47230-3. In French.
- [6] H. Petitmangin. *Grammaire latine*. Nathan, Paris, revised ed., 2018. ISBN 978-2-09-171001-3. In French.
- [7] E. Plantade. Aspects métriques et rythmiques de la couleur archaïque dans les distiques élégiaques d’Apulée. In *Stylistique et poétique de l’épigramme latine : Nouvelles études*, D. Valat, F. Garambois-Vasquez, eds., pp. 139–153. MOM Éditions, Lyon, 2022. <https://doi.org/10.4000/books.momeditions.16900>
- [8] L. Sausy. *Grammaire latine complète*. Eyrolles, Paris, 8th ed., 2019. ISBN 978-2-212-54685-9. In French.

(Since this article focuses on typography issues, the publication years of the books mentioned in the bibliography correspond to the date of the *dépôt légal* (“registration of copyright”) so as to avoid ambiguity regarding the considered printing.)

A User commands and settings

We give the user-defined commands and settings used hereinbefore. (The author has partially relied on information found on tex.stackexchange.com.)

First, to typeset a breve and a macron over diphthongs, in both \LaTeX and Lua \LaTeX :

```
\newcommand*{\dbreve}[1]{\breve{\hbox{#1}}\m@th$}
\newcommand*{\dmacron}[1]{\overline{\hbox{#1}}\m@th$}
\newcommand*{\dmacronkern}[1]{\mkern 1.5mu\overline{\mke
```

(The `\dmacronkern` command produces a slightly shortened macron, as demonstrated.)

Second, to adjust the vertical gap, for instance to 1.6 pt, between a macron (obtained with `\dmacron` or `\dmacronkern`) and letters, in Lua \LaTeX :

```
\check@mathfonts
\Umathoverbarvgap\textstyle=1.6pt
```

(Note that this setting is declared within the main document, that is, not in the preamble, and needs to appear *after* the `\maketitle` command, if ever called.)

Third, negative kerning, for instance of -1.5 pt, around a vertical bar can be obtained as follows:

```
\kern-1.5pt\textbar\kern-1.5pt
```

Finally, when denoting elision with a stretched breve below the baseline and there is an elision at the cæsura, we have used a `\makebox` of zero width to typeset a double bar over an undertie:

```
□\symbol{"203F}\makebox[0pt]{\textbardbl}□
```

This could be considered future work: the stretched breve covers two characters, but three would be better in this case (i.e. a double bar surrounded by spaces).

B Related Unicode glyphs

A summary of the Unicode glyphs mentioned in this note mostly for diacritical marks and ties is given in the table below.

Description	Code point
Breve	U+02D8
Combining overline	U+0305
Combining breve	U+0306
Combining double breve below	U+035C
Combining double breve	U+035D
Combining double macron	U+035E
Undertie	U+203F
Inverted undertie	U+2054

- ◇ Antoine Bossard
Kanagawa University, Graduate
School of Science, Yokohama
221-8686, Japan
abossard (at) kanagawa-u.ac.jp
ORCID 0000-0001-9381-9346

Ac dubitabam tunc an non Pragam essem aditurus.
 Āc dūbī|tābām | tūnc || ān | nōn Prālgam_ēssem_ādī|tūrūs.