

A short note on typesetting Latin verse scansion with LaTeX and LuaLaTeX

ANTOINE BOSSARD, KANAGAWA UNIVERSITY (JAPAN)

TUG 2024, PRAGUE, JULY 19-21

Latin verses

Prosody, metrical feet

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.

Large parts of the Latin literature are written in verse:

- Virgil's *Æneid*
- Ovid's *Metamorphoses*
- etc.

Scansion

Feet, cæsura

Scansion of a Latin verse is about identifying **feet**.

This requires evaluating the quantity of the verse syllables.

To this end, **diacritical marks** are used:

- a breve (˘) denotes a short vowel,
- a macron (¯) a long one, and
- both diacritics for a common syllable.

A verse is usually split into two with a **cæsura**.

While feet are conventionally separated by vertical bars (|), the cæsura is often materialized with a double vertical bar (||).

Other symbols are used for **elision** and **hiatus**.

The case of LaTeX

Single and multiple diacritics

No issue when dealing with a single diacritical mark for one character.

Both accented Unicode characters (e.g. U+0103 `ǎ`, U+0100 `Ā`) and letters manually marked with the commands `¥u` (breve), `¥=` (macron) (e.g. `¥u{a}`, `¥=A`) are correctly rendered.

On the other hand, the stacking of several diacritical marks atop one letter remains difficult.

For instance, `¥u{¥=a}` is of no avail as it produces `˘ā`.

Several packages are here to help: `tipa`, `stackengine`, etc.

The case of LaTeX

Diphthongs

The rendering of `æ` and `œ` is acceptable: `ǣ`, `Ǣ`.

Yet, the duration of the `æ` diphthong can be short: the corresponding characters are rendered as expected.

It is however more difficult to add a diacritical mark above the other diphthongs, such as `au` and `eu`.

Relying on the `¥overline` command is one solution. But the bar is this time slightly too long and is, more or less depending on the font, too close to letters.

While the former issue can be addressed with `¥mkern`, the latter remains.

The case of LuaLaTeX

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.

Assuming a font that is aware of such diacritical issues has been selected (e.g. Noto), LuaLaTeX brings significant advantages over LaTeX as far as we are concerned:

- It enables **multiple diacritical marks** on a character.
- The **macron** over diphthongs can be further improved: width and vertical gap can be **adjusted**.

About Unicode

Unicode does provide glyphs that could be helpful for our purpose. We did some experiments with Noto Serif:

1. The **Combining overline** glyph (U+0305), whose purpose is to render a continuous line above several letters.
Unsatisfactory: the two overlines are not combined at all.
2. The **Combining double macron** glyph (U+035E).
Unsatisfactory: although spanning both letters, the macron is oddly positioned and too short.
3. The **Combining double breve** glyph (U+035D), whose purpose is to add a breve above two letters. **Satisfactory**.
4. And a few others (refer to the TUG article).

Summary

Regarding [LaTeX](#), the support of single diacritical marks applied to a single letter is fully satisfactory.

For “native” diphthongs such as æ and œ, macrons and breves are rendered as expected. Regarding other diphthongs, like au and eu, simple commands could do but results vary.

The most problematic issue arises from double diacritical marks: their support requires an additional package.

Regarding [LuaLaTeX](#), Unicode provides useful features for typesetting Latin verse scansion. However, their support strongly depends on fonts.

Epilogue

Sample Latin verses with scansion

ille pedum melior motu fretusque iuventa,
īllē pē|dūm mēlī|ōr || mō|tū frē|tūsquē jū|vētā,

hic membris et mole valens; sed tarda trementi
hīc mēm|brīs ēt | mōlē vā|lēns; || sēd | tārđā trē|mēntī

genua labant, vastos quatit aeger anhelitus artus.
gēnūā lā|bānt, vās|tōs || quātīt | āgēr ān|hēlītūs | ārtūs.

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

(Virgil, *Æneid*, book V, lines 430–433)