Accent in Latin Hexameter Poetry

Ben Nagy

benjamin.nagy@adelaide.edu.au

May 28, 2019

As part of my statistical analysis of hexameter poetry, I was interested in analysing the interplay between metrical ictus and word accent. To this end, I wrote a short program that was able to automatically analyse pre-scanned text to determine the conflicts. To do this, it was first necessary to be able to correctly place the accent on any Latin word. While this seems simple, it was more complicated than it may appear. This short research summary aims to record and justify my approach.

Keywords: Latin, digital humanities, python

1 INTRODUCTION

I have no choice but to assume, in this, a reasonable level of familiarity with the particular jargon of Latin poetry. I should, however, clarify how I am using the key terms. An *ictus syllable* is one which begins a foot, and in dactylic hexameter is always long (this might also be referred to as the *arsis*). The *accent* of a Latin word, which some might call *stress* is extra emphasis placed on a syllable. For our present purposes, I am not concerned at all with the distinction between a pitch accent or a stress accent, nor indeed the potential technical difference between the terms *stress* and *accent*, and will use them interchangeably. It suffices to know that multi-syllable Latin words bore one (and only one) primary accent. The rule for determining where the accent falls in a Latin word is, ostensibly, straightforward. In words of two syllables, the first is always stressed.[1] For words of more than two syllables, according to the 'Law of the Penult', the *penult* (second-to-last syllable) is stressed if long, and the *antepenult* (third-to-last) otherwise. The rule, however, is not as universal as one might think. To conduct this research, it was necessary to automatically determine where, in a line of dactylic hexameter, the ictus fell in each foot, and whether or not that syllable was also stressed. It should, therefore, be no surprise that, in the course of making those determinations for tens of thousands of lines, there were some edge cases. For those edge cases, the algorithm had to be able to make an automatic decision. This section summarises the research that informed the final design of the algorithm.

---

1 This is, as we will see, a lie.

The website Musisque Deoque (2007) has a collection of most of the poetry in the Classical Latin corpus, which includes its metrical scansion. Those texts have now been made available for download in XML format, suitable for automatic analysis. All of my analysis is based on the MQDQ (XML) editions of the relevant works. In general, each line of hexameter was treated individually. The basic scansion of the lines is part of the MQDQ data format. Each word in a line is tagged with the syllables it contains, as well as a variety of extra metrical features. In particular, the MQDQ data marks elision (or *synalepha*) diaeresis, prodelision and strong and weak caesurae. The syllables that make up a word are marked according to the feet in which they fall, and whether they are in the arsis or thesis—so 3T is the thesis of the third foot.

```
<line name="1" metre="H" pattern="DDSS">
    <word sy="1A1b" wb="CF">Arma</word>
    <word sy="1c2A2b" wb="CF">uirumque</word>
    <word sy="2c3A" wb="CM">cano,</word>
    <word sy="3T4A" wb="CM">Troiae</word>
    <word sy="4T" wb="DI">qui</word>
    <word sy="5A5b" wb="CF">primus</word>
    <word sy="5c" wb="DI">ab</word>
    <word sy="6A6X">oris</word>
</line>
```

Figure 1: Raw XML for a line from MQDQ

The calculation of word accent is not part of the MQDQ data, and was performed by my `mqdq-parser` code (a work-in-progress, open-source Python library available at `https://github.com/bnagy/mqdq-parser` (Nagy (2019))).

2.1   *Feet or Words?*

The first choice was whether to calculate conflict by foot, or by word. Based on the analysis in Knight (1931), where he pays special attention to the fourth foot, it seems that traditional analysis has been based around feet (not words). This is more logical - every foot must have an ictus, and that syllable is either accented or it is not. If the analysis were carried out with respect to words, then a given word might not contain an ictus and so the metric would not apply. Classifying the harmony by foot means that, in every line, we have exactly six places where ictus and accent may (or may not) coincide. Knight calls feet in which the ictus and accent fall on the same syllable *homodyne*, and the feet in which they do not *heterodyne*. Although this terminology is not particularly satisfactory, it has the benefit of being compact, and I will adopt it from this point forward.

2.2 Handling Clitics

The first question that arises is how to handle words with clitics (-que, -ve, -ne etc.). For example, *cǎ.nō* is stressed on the first syllable (according to the Law of the Penult), but what about *cǎ.nō.quě*? This is covered in Allen's landmark *Vox Latina* (Allen, 1965, 87):

> When an enclitic (*-que, -ue, -ne, -ce*) was added to a main word, the resulting combination formed a new word-like group, and a shift of accent was therefore to be expected in some cases: thus, for example, *vírum* but *virúmque* (like *relínquo*). Such a shift is discussed by many of the grammarians [. . .]

Allen then refers us to the excellent work by Schoell (1876), *De Accentu Linguae Latinae*, in which "the majority of the ancient observations on the Latin accent are collated". However, as he discusses, the grammarians seem to have over-zealously created a new rule whereby the accent *always* moved to the penult in the presence of an enclitic. This, he notes, this does not seem to be correct, "supported by the fact that such combinations are commonly found in the fifth foot of hexameters, where we expect agreement between ictus and accent" (Allen, 1965, 87). How commonly? I find 1037 instances in Verg. *Aen.* of a word in the fifth foot, ending with *-que*, where the arsis falls on a long penult (*pūgnā́squě, sācrū́mquě* etc). If we did not already know it, this would certainly convince us of the fact that the *-que* clitic causes the word accent to move. However, I also find 111 instance of words in the fifth foot like *āltǎquě*, *āctǎquě, ōssǎquě* which have a light penult. While I am not sure that 111 lines from ~10,000 comprise a 'common' pattern, I am inclined to agree with Allen that the accent here is very likely to fall on the antepenult (*ā́ltǎquě*), not the penult—*pace* Varro, Capella *et cetera*. With these assumptions, I find only 127 lines (of 9840 uncorrupted lines in the MQDQ edition) with a heterodyne fifth foot. Interestingly, many of them end with five syllable words like *Pirithoumque* or *quadrupedantum* which begin on the arsis of the fifth foot. Linguistically, there are good reasons to suspect that long words like these would carry a secondary stress, which would naturally fall on the first syllable, making these fifth and sixth feet 'almost' homodyne (*quǎ́drupedántum*). This could be an interesting addition to the arguments in favour of a secondary word stress.

2.3 Handling Elision

While the grammarians collected by Schoell are full of advice about the pronunciation of *póne* versus *poné*,[2] they are entirely silent on the matter of elision. Of the several hundred excerpts, I could not find a single one that considered the question of whether elision, by changing the number of syllables in a word, moves the accent. This, in itself, seems to be evidence for the fact that nothing happens—in other words, that the accent does not move when elision occurs. The situation is made more complicated, however, by the fact that we can't simply look at some examples and "figure it out". Consider, for example, this line from Verg. *Aen* (3.658):

---

2 This (potential) difference in pronunciation is mentioned in no less than twenty passages.

$$— \quad —\mid— \quad —\mid— \quad —\mid— \quad —\mid— \; \smile \; \smile\mid— \quad \times$$

Monstrum horrendum informe ingens cui lumen ademptum

Vergil uses this singular metre, with the drag and thud of three spondaic elisions, to add an extra dimension to his description of Polyphemus. In such an unusual hexameter line, it is not, I claim, entirely obvious where the words should be stressed. Buoyed by the success of the previous approach (examining the fifth feet in hexameters), I hoped that we could do something similar, but elisions in the fifth foot are rare. Of the 9840 lines of the *Aeneid*, there are 132 lines where the word containing the arsis of the fifth foot has an elision. Ovid is even more particular—in 12009 lines of the *Metamorphoses* we find just 66. With so little data, it seems dangerous to draw conclusions—could these lines not simply be heterodyne in the fifth foot? After all, we know that that 'rule' is not universally applied. Nevertheless, by manual inspection, there do seem to be patterns in Vergil which suggest that he intended most of them to harmonise. A few examples should suffice. If we make the assumption that the accent does not move with elision, there are several lines which would end with a 'normal' bucolic diaeresis: *. . . crímin(e) ab úno* (*Aen.* 2.65), *. . . sánguin(e) et ígni* (*Aen.* 2.210), *. . . lítor(e) ab ómni* (*Aen.* 5.43). However, another 'rule' seems to be that if a clitic is elided, the accent returns to its normal position. This is seen in several lines with a sixth foot disyllable and a (presumably) stressed fifth foot arsis: *. . . fatáliaqu(e) árva* (*Aen.* 5.82), *. . . spumántiaqu(e) áddit* (*Aen.* 5.187) *. . . ábditaqu(e) íntus* (*Aen.* 9.579). These cases, though, all involved elision after breves. Elision after a long syllable is even rarer. For that, I decided, in the end, to leave the stress where it was; but I acknowledge the paucity of data. There are only four lines of the *Aeneid* with an elision after the arsis of the fifth foot that do *not* involve *-que*. One is *Aen.* 3.164:

$$— \; \smile \; \smile \mid — \; \smile \; \smile\mid— \quad —\mid— \quad \smile \; \smile\mid— \quad \smile \; \smile\mid— \quad \times$$

Nomine Achaemenides, Troiam genitore Adamasto

If this line is one of the extremely rare[3] occurrences of fifth foot heterodyne, I can find no poetic reason for it, other than the minor difficulty of wrangling 'Achaemenides'—surely routine for a poet like Vergil. I choose instead to believe that *genitór(e)* retains its stress on what has become its final syllable. With these slight clues, and based on the absence of any discussion by the ancient grammarians, I implemented elision with two simple rules: if a word with elision ends with a long syllable, that syllable is stressed. If it ends with a breve, then the word is stressed exactly as if it were a 'normal' word according to the remaining syllables. In particular, an elided clitic is not restored when determining stress. As a very final point, if my approach finds no support (nor any disagreement) in the ancient sources, I seem at the least to be in accord with (Westaway, 1930, 68), whom (Knight, 1931, 186 n. 1) cites as an authority:

> [under 'Subsidiary Rules']
>   (2) Sometimes a final syllable following a long penult is lost, or there is a contraction. In such cases the accent is retained on the syllable which has become final.

---

3 As noted above, there are 127 in Verg. *Aen.* although reasoning by this count is slightly circular logic.

2.4 Exceptions and Special Cases

There are several remaining exceptions and special cases notes by (Allen, 1965, 87), some of which I have implemented, and some not. The first two classes are these:

> In some words, originally accented on the penultimate, the vowel of the final syllable has been lost; and the accent then remains on what is now the final syllable. Thus, for example, *nostrās, illīc, adhūc, addūc, tantōn* (from *nostrātis, illīce, adhūce, addūce, tantōne*). The same applies to contracted perfect forms in *-āt, -īt*, from *-āvit, -īuit* [. . .]

The first group of words were easy enough to add as special cases, although it should be noted that they are extremely rare in Verg. *Aen.* (*nostras* occurs just four times). The contracted perfect forms, however, were not considered. To be able to detect these forms, as opposed to the many, much more common, forms that end with *-at* or *-it*, would require an automatic parser with a level of accuracy and sophistication that does not yet exist.

The final group of exceptions that were implemented are a small set of conjunctions. Allen notes merely that "certain conjunctions were unaccented, e.g. *at, et, sed, igitur*" (Allen, 1965, 88), but refers us to Schoell for the details. I followed the broad consensus, summed up in this excerpt from *Mus. Rhen.* 8 (Schoell, 1876, 194):

> Copulativae et disiunctivae prope omnes gravantur, expletivae plures fastigia retinent, causales autem et rationales quaedam cum fastigiis, alias gravi accentu deprimuntur. quaedam in pronuntione deprehendes.

> Copulative and disjunctive [conjunctions] are almost all lowered [unaccented], most expletives retain their fastigia [accent marks], however some causals and conclusives are with fastigia, others are lowered by a grave accent. Some you will discover as you pronounce them.

With this in mind, I produced the following list, comprising conjunctions, disjunctions, *igitur* (from Allen) and some others of one syllable: *at, ac, atque, et, sed, igitur, vel, aut, iam, seu, nam*. In some cases, polysyllabic conjunctions are suggested by the grammarians (*quoniam, etenim, sedenim*), but I have not marked them as unaccented. It seems to me, in those cases, wiser to err on the side of caution. After all, we have already seen from our exploration of *-que* that the early grammarians are not fully authoritative. The final category that should be mentioned are a group of prepositions which, per Allen and Schoell, may have changed their accentuation depending on whether they appeared before or after the relevant noun. Once again, detecting these automatically would be beyond our present level of parsing technology.

2.5 *Replicating Knight*

It is difficult to convincingly assess the 'accuracy' of my final accentuation algorithm. Fortunately, however, there is at least one point of comparison—in Knight's *Homodyne in the fourth foot of the Vergilian Hexameter*. He informs us that "[f]ourth-foot homodyne is by comparison rare in Vergil, but not very rare. The percentage of it to the total number of lines in the *Aeneid* is 35.95 per cent" (Knight, 1931, 186). Using my algorithm, I detect fourth-foot homodyne in 3507 lines (of 9840), which is 35.65%. The difference in our percentages represents about 30 lines on which we disagree. Knight also lists some percentages by individual books. For *Aen.* 1, he has 28.61% while I have 28.82%; *Aen.* 8, 39.97% to my 39.29%; *Aen.* 10, 32.59% vs 32.71%. While Knight does not share with us his methodology for scanning and tabulating the entire *Aeneid* in the 1930s (perhaps it was a long winter), he does mention that he performed "two computations, in which different texts were used and different conventions followed" and that his figures varied by only 0.45% (Knight, 1931, 186 n. 1). Based on this, it seems reasonable to trust my own figures, certainly for performing bulk statistical analysis, although it might be wise to verify the calculation for small passages by hand.

REFERENCES

Allen, W. S. (1965). *Vox Latina*. Cambridge: CUP.

Knight, W. F. J. (1931). Homodyne in the fourth foot of the Vergilian hexameter. *The Classical Quarterly*, 25(3-4), 184–194.

Musisque Deoque. (2007). *Musisque Deoque (MQDQ) A digital archive of Latin poetry.* Retrieved May 28, 2019, from `http://mizar.unive.it/mqdq/public/`

Nagy, B. (2019). *mqdq-parser - utility scripts in Python for working with poems from MQDQ.* Retrieved May 28, 2019, from `https://github.com/bnagy/mqdq-parser`

Schoell, F. (1876). De accentu linguae Latinae. In F. Ritschelius (Ed.), *Acta societatis philologae lipsiensis,* VI (pp. 1–231). Leipzig: Teubner.

Westaway, F. W. (1930). *Quantity and accent in Latin*. Cambridge: CUP.