Metrical tension and prose cadence

Liam Smith and Arto Anttila (Stanford University)

Overview. Across poetic traditions, lines of verse are known to exhibit METRICAL CLOSURE: beginnings are free, endings are strict (Smith 1968, Hayes 1989, Ryan 2017, Kiparsky 2020). Fabb (1997) calls the freer first half of the line the 'free initial' and the rigid second half the 'fixed cadence'. Inversion is a showcase example: *Making a famine where abundance lies* is a line of iambic pentameter that starts with a trochaic word (*Máking*) while the rest of the line perfectly aligns strong and weak positions with stressed and unstressed syllables. But is metrical closure a property of verse or is it a more general property of language? Does the prose sentence *Soap is ordinarily thought of as the common cleansing agent well known to everyone* have a free initial and a fixed cadence? The question is difficult to answer because prose is unmetered and measuring metrical well-formedness presupposes a scansion. Fortunately, it is possible to make progress on this question by using the Prosodic method (Heuser et al. 2010) that allows one to automatically scan any text, verse or prose, using identical metrical constraints without presupposing the existence of an actual scansion. We present new evidence for fixed cadences in prose and conclude that metrical closure is an inherent property of ordinary language that is simply highlighted in metrical verse.

Theory and method. We assume the metrical theory defined in Hanson and Kiparsky 1996 and implemented in Prosodic (Heuser et al. 2010) that builds on Optimality Theory (Prince and Smolensky 1993/2004). Given a raw input text, Prosodic annotates it for phonological variables, including stress and syllabification. Among the CANDIDATE SCANSIONS, i.e., possible sequences of strong and weak positions (e.g., *wswswswsws, wwswwswwsw*, etc.) paired with input syllables, Prosodic finds the VIABLE SCANSIONS defined as those scansions not harmonically bounded by the constraint set. We adopted five metrical constraints (Hanson and Kiparsky 1996, de la Fuente et al. 2023): *W/PEAK: A weak position must not contain a strong syllable (= stressed syllable of a polysyllabic word); *S/UNSTRESSED: A strong position must not contain an unstressed syllable; *W/STRESSED: A weak position must not contain a stressed syllable; W-RESOLUTION: For disyllabic positions within a word, the first position must be light and stressed (*many/*although*); F-RESOLUTION: A disyllabic position across a word boundary must be weak with two function words (*as the/*light's flame*). We assumed that resolution maximally allows two syllables per position. As a measure of metrical well-formedness, we used METRICAL TENSION SUM (MTS) defined as the sum of violations of all the viable scansions of the line divided by the number of syllables in the line (Anttila et al. 2022). For example, the initial substring *Making a famine where abundance* permits 4 viable scansions with a MTS of 10, whereas the final substring *famine where abundance lies* has 1 viable scansion and a MTS of 0. For our tests, we used the current version of Prosodic (`https://github.com/quadrismegistus/prosodic`), adopting the constraint definitions as they stand. We note that violations of F-RESOLUTION seem doubled, but we believe that is unlikely to affect the results.

Data. We constructed a dataset of texts from 51 different sources, collected from Project Gutenberg and The American Presidency Project (Woolley and Peters 2008), which span across four genres: Verse, Speeches, Procedural Texts, and Theatrical Scripts. From these texts, we extracted a total of 28,403 lines. For verse, the original lineation was respected. For all other genres, we lineated at full stops. Lines less than 8 syllables long or which included non-English characters were excluded. We then processed the data into a parallel dataset of both the first and last ten syllables of each line. To extract the initial and final ten syllables, we used the *Syllables* library (Day 2023) in Python to take ten syllables from the beginning/end of the line, respecting word boundaries.

We then parsed each line using the *Prosodic* library (Heuser et al. 2010), which outputs all viable scansions for a line of text and which metrical constraints each scansion violates. The scansions for each line were then collapsed into a single entry including aggregate data across all scansions and the data from the best scansion.

Results. We first verified that MTS behaves as expected across genres. MTS is lowest in Verse, higher in spoken genres (Theatrical Scripts, Speeches), and highest in written genres (Procedural Texts). This is in general agreement with earlier studies (Anttila and Heuser 2016). We also found that metrical heterogeneity operationalized as the variance of MTS across lines increases in the same

order: verse (mean = 0.65, variance = 1.23), theatrical scripts (mean = 1.11, variance = 2.48), speeches (mean = 1.41, variance = 3.68), procedural texts (mean = 1.53, variance = 4.43). We speculate that these differences among genres reflect their textual function. Texts intended to be spoken or heard tend to engage more with phonology and meter, which in turn results in stricter stress-meter alignment and lower metrical tension. Informational texts that primarily exist in the written mode can more easily ignore phonology. This results in looser stress-meter alignment and higher metrical tension.

We tested the closure hypothesis (beginnings are free, endings are strict) using linear regression. Crucially, we included word length as a control variable. It has been reported that prose favors longer words than verse (Anttila and Heuser 2016). This may reflect differences in topic but also the fact that long words are harder to employ in verse because of metrical constraints. Genres also seem to differ in how constituent length matters to linearization: in utility prose the tendency is long last ("end weight") whereas in verse and art prose the opposite tendency has been observed (Saintsbury 1912, Croll 1919, Blumenfeld 2016). This motivates the inclusion of word length as a predictor.

Our simplest model `lm(tension ~ initial + mean_word_length)` involves minimal arbitrary decisions on our part: the response variable is the MTS calculated by Prosodic; the predictors line-initiality and mean word length are mechanically calculated from the text. The closure hypothesis turned out to be supported: MTS is positively associated with line-initiality ($b = 0.11$, $SE = 0.01$, $p < 0.001$) and mean word length ($b = 0.98$, $SE = 0.018$, $p < 0.001$). In other words, MTS is higher in line-initial strings and increases with word length, just as expected. We then added the fixed effect `genre` with four levels (Verse, Theatrical Scripts, Speeches, Procedural Texts) using Helmert-coding. The initiality and word length effects persisted ($p < 0.001$). To control for more fine-grained differences among individual source texts we fitted a mixed effects linear regression model with initiality and mean word length as fixed effects, adding `source` (51 groups) as a random intercept. Again, the effects persisted ($p < 0.001$). Interestingly, the metrical closure effect only emerges if we control for word length; it only becomes visible once the masking effect of word length is removed. This may be the reason why it has not been observed earlier. Finally, one would like to understand why sentences should begin metrically lax and become stricter towards the end. Following Ryan's (2017) suggestion, we speculate that meter is more strictly regulated in syllables under higher levels of phrasal stress which in English typically piles up towards the end of the sentence reflecting the right-branching syntax of the language accompanied by cyclic phrasal stress (Chomsky and Halle 1968).

Conclusion. We conclude that metrical closure is not a property of metered verse but a general prosodic property of ordinary language.