Metrical uncertainty

Proposal. Halle and Keyser (1966, 1971) observed that experienced readers are capable of distinguishing, not only metrical lines from unmetrical lines, but also more complex lines from less complex lines. This raises the question of how to characterize metrical complexity in a way that correlates with human judgments, reliably distinguishes different types of verse, and generalizes from verse to prose. Here we test a simple measure of metrical complexity, METRICAL UNCERTAINTY (Anttila et al. 2022), and show it captures a subtle difference between two types of verse (Hanson 2006) and a less subtle difference between verse and prose.

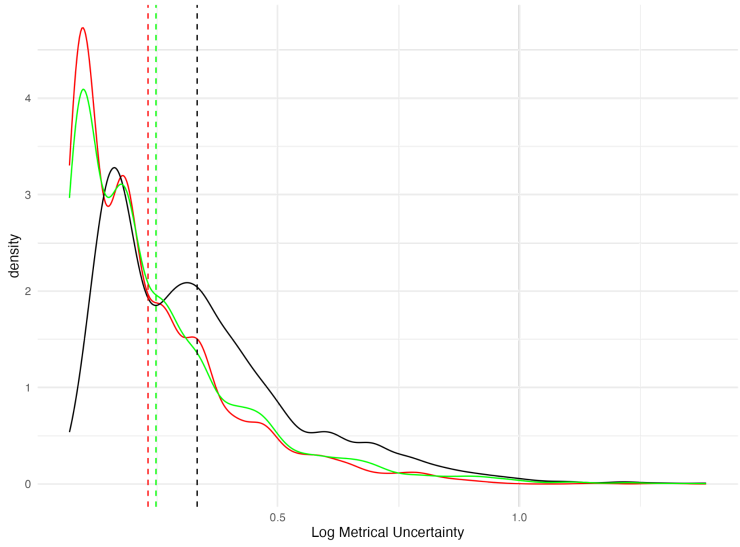
Metrical analysis. We define SCANSION as a correspondence relation between abstract verse templates and linguistically structured texts (Kiparsky 2020). We assume that metrical correspondence is governed by the following VIOLABLE CONSTRAINTS that play a central role in English verse (Hanson and Kiparsky 1996): *W/PEAK 'A weak position must not contain a strong syllable', *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', and F-RESOLUTION 'A disyllabic position crossing a word boundary must be weak and its two words must be function words'. To do metrical scansion automatically, we used PROSODIC (Heuser et al. 2010-) that identifies the VIABLE SCANSIONS of a line from among the logically possible scansions (for a 10-syllable line, $2^{10} = 1{,}024$) by limiting resolution to at most two syllables and discarding all harmonically bounded scansions (Prince and Smolensky 1993/2004). This leaves a (typically small) set of viable scansions, each one a possible winner depending on constraint ranking. The metrical uncertainty of a line is defined as the number of viable metrical scansions. Intuitively, the more scansions compete for the reader's/hearer's attention, the more metrically complex the line, and selecting the correct scansion may be difficult or even impossible, as is often the case in prose. Metrical uncertainty allows us to quantify the metrical distance between different kinds of verse and since the method is fully automatic it can also be applied to prose.

Shakespeare's lyric and dramatic metrical styles. Hanson (2006) pointed out a metrical difference between Shakespeare's metrical styles in *The Sonnets* (lyrical style) and *Richard II* (dramatic style): the dramatic style is looser and "inclusive of the lyrical style", suggesting that the metrical mappings allowed in *The Sonnets* are a subset of those allowed in *Richard II*. We scanned the 2,155 lines of *The Sonnets* and 2,800 lines from *Richard II* using Prosodic and measured the metrical uncertainty of each line. A logistic regression model revealed that the number of viable scansions (normalized by the number of syllables in the line) is a good predictor of metrical style: an increase in the number of scansions increases the probability of the line coming from *Richard II* (logit difference 0.38, $SE = 0.11$, $z = 3.41$, $p = 0.0006$), confirming Hanson's result, and showing that metrical uncertainty is able to detect subtle differences between metrical styles. We also calculated the overall variance of the number of parses for all of *The Sonnets* (= 5.95) and all of *Richard II* (= 8.57). In addition, we took the variance within each individual Sonnet (excluding Sonnets 99 and 126 which do not have 14 lines), as well as 14 line chunks from a 2,155 line subset of *Richard II*. In accordance with Hanson's hypothesis, the minimum variance of all the 14 line chunks is not dissimilar between *Richard II* (0.55) and *The Sonnets* (0.42) but the maximum variance is much higher for *Richard II* (64.41) than for *The Sonnets* (35.64). These results imply that our metrical uncertainty measure can capture subtle stylistic differences independently established.

Comparing verse and prose. Metrical uncertainty also separates verse from prose. Based on data from Project Gutenberg (manuals) and The American Presidency Project (Peters and Woolley 1999-), a logistic regression model revealed that the (normalized) number of viable scansions is a good predictor of the prose/verse difference: a decrease in the number of scansions increases the probability of the line being verse (logit difference -1.39, $SE = 0.10$, $z = -14.44$, $p < 2e\text{-}16$), showing that metrical uncertainty is able to detect the metrical difference between verse (*The Sonnets, Richard II*) and prose. The density plot below shows the metrical uncertainty of lines across the three text types: the red line is *The Sonnets*, the green line is *Richard II*, the black line is prose. The dashed lines are the means.

Future work. Metrical uncertainty as defined above assumes that all constraint violations are equally serious. Two obvious extensions are METRICAL TENSION SUM (Anttila et al. 2022) that sums up the number of violations across all viable scansions and WEIGHTED METRICAL TENSION SUM that allows constraints to be independently weighted as in HG/MaxEnt (Hayes, Wilson and Shisko 2012, Hayes 2017). These enriched measures are likely to allow a more fine-grained metrical separation of text types.

Summary. We have tested a simple measure of metricality, METRICAL UNCERTAINTY, that turns out to be able to separate Shakespeare's lyric verse (*The Sonnets*) from Shakespeare's dramatic verse (*Richard II*), and both of them from prose. This provides an operationalized alternative to earlier metrics, in particular Halle and Keyser's (1971) notion of metrical tension.



Selected references: Anttila, Arto, Ryan Heuser & Paul Kiparsky. 2022. Prose rhythm and antimetricality. *Literary Linguistic Forms*, 2022 LSA Annual Meeting; Halle, Morris & Samuel Jay Keyser 1971. *English stress: Its form, Its Growth, and Its Role in Verse*. New York, Evanston & London: Harper & Row, Publishers; Hanson, Kristin. 2006. Shakespeare's lyric and dramatic metrical styles. In *Formal approaches to poetry*, eds. Bezalel Elan Dresher and Nila Friedberg, 111–133, Berlin: Mouton de Gruyter; Hanson, Kristin and Paul Kiparsky. 1996. A parametric theory of poetic meter. *Language* 72: 287– 335; Heuser, Ryan, Joshua Falk & Arto Anttila. 2017. Prosodic v0.2, a linguistic annotator and metrical parser, software package, Stanford University. https://github.com/quadrismegistus/prosodic.