Metrical tension and prose cadence

Liam Smith and Arto Anttila

Stanford University

AMP 2024

Rutgers University
November 2, 2024

While some lines of poems start loose
They tighten up towards the caboose
This abstract suggests
That non-metrical texts
Show similar effects, though reduced.

(an anonymous AMP reviewer)

Overview

- Across poetic traditions, lines of metrical verse exhibit METRICAL CLOSURE: beginnings are lax, endings are strict.

  (Smith 1968; Hayes 1983; Fabb 1997; Ryan 2017; Kiparsky 2020).

- We provide evidence that metrical closure is not only a property of metered verse but a general property of language.

- Gradient metrical markedness parallels gradient phonological markedness (e.g., hiatus avoidance) that has been observed in sentence formation (Breiss and Hayes 2020).

Scansion in metrical verse

- Scansion is a correspondence relation between metrical positions (w = weak, s = strong) and syllables.

  (Halle and Keyser 1971; Kiparsky 1977; Prince 1989; Hayes et al. 2012; Blumenfeld 2015, a.o.).

- A well-behaved scansion (Shakespeare, Sonnet 2, 1):

w	s	w	s	w	s	w	s	w	s
When	fór-	ty	wín-	ters	sháll	be-	síege	thy	bród

- This scansion perfectly satisfies the following two constraints:
  - *w/stressed: No stressed syllable in a weak position
  - *s/unstressed: No unstressed syllable in a strong position

Scansion in metrical verse

- INVERSION is common line-initially and after mid-line pauses (beginnings are lax); see, e.g., Steele 1999.
- An example of inversion (Shakespeare, Sonnet 1, 7):

w	s	w	s	w	s	w	s	w	s
Má-	king	a	fá-	mine	whére	a-	bún-	dance	líes

- *Máking* violates both *w/STRESSED and *s/UNSTRESSED.

Scansion in prose?

- Does prose exhibit metrical closure?

  Soap is ordinarily thought of as the common cleansing agent well known to everyone.
  (E. G. Thomssen, *Soap-Making Manual*)

- It is hard to tell: prose is unmetered and identifying metrical violations presupposes a scansion.

The Prosodic Method (Heuser et al. 2010):

Heuser, Ryan, Joshua Falk, and Arto Anttila. 2010–.
Prosodic (software), Stanford University.
`http://prosodic.stanford.edu/`
`https://github.com/quadrismegistus/prosodic`

- allows the automatic scansion of any text
- does not presuppose an actual scansion
- applies the same metrical constraints to prose and verse
- makes prose and verse metrically comparable

The Prosodic Method (Heuser et al. 2010):

- Theoretical assumptions:
  - metrical constraints (ours are from Hanson and Kiparsky 1996)
  - embedded in Optimality Theory (Prince and Smolensky 1993/2004)
- Prosodic takes the following steps:
  - analyses input phonologically (e.g., word stress, syllabification)
  - generates CANDIDATE SCANSIONS
  - assigns metrical constraint violations to candidate scansions
  - identifies VIABLE SCANSIONS, i.e., those not harmonically bounded

Three candidate scansions (out of $2^{10} = 1,024$)

- Candidate scansion 1 (iambic pentameter):

w	s	w	s	w	s	w	s	w	s
When	fór-	ty	wín-	ters	sháll	be-	síege	thy	bród

- Candidate scansion 2 (trochaic pentameter):

s	w	s	w	s	w	s	w	s	w
When	fór-	ty	wín-	ters	sháll	be-	síege	thy	bród

- Candidate scansion 3 (trochaic tetrameter with resolution):

s	w	w	s	s	w	w	s	w	s
When	fór-	ty	wín-	ters	sháll	be-	síege	thy	bród

Metrical constraints (Hanson and Kiparsky 1996)

- *S/UNSTRESSED ("THE HOPKINS CONSTRAINT")
  A strong position must not contain an unstressed syllable.

- *W/STRESSED
  A weak position must not contain a stressed syllable.

- *W/PEAK ("THE SHAKESPEARE CONSTRAINT")
  A weak position must not contain a stressed syllable of a polysyllabic word (e.g., *póison*).

- W-RESOLUTION
  For disyllabic positions within a word, the first position must be light and stressed (e.g., *mány*/**althóugh*).

- F-RESOLUTION
  A disyllabic position across a word boundary must be weak with two function words (e.g., *as the*/**líght's fláme*)

Measures of metrical well-formedness

- Metrical Uncertainty (MU)
  The number of viable scansions divided by the number of syllables in the line (de la Fuente, Nick, and Anttila 2023).
- Metrical Tension Sum (MTS)
  The sum of violations across all viable scansions divided by the number of syllables in the line (Anttila, Heuser, and Kiparsky 2022).

Sample line fragment	MU	MTS
Making a famine where abundance	4	10
famine where abundance lies	1	0

- Resolution: Maximally two syllables per position (`s` or `w`).

Data

- *Project Gutenberg*, *The American Presidency Project* (Woolley and Peters 1999–)
- 51 texts; 4 genres; 28,403 lines
- Lineation in different genres:
  - For Verse, original lineation was respected.
  - For Speeches, Theatrical Scripts, and Procedural Texts we lineated at full stops.
- Two parallel datasets:
  - 10 line-initial syllables
  - 10 line-final syllables
- Excluded: lines with < 8 syllables or non-English characters
- Extracted using *Syllables* (Day 2023) respecting word boundaries.

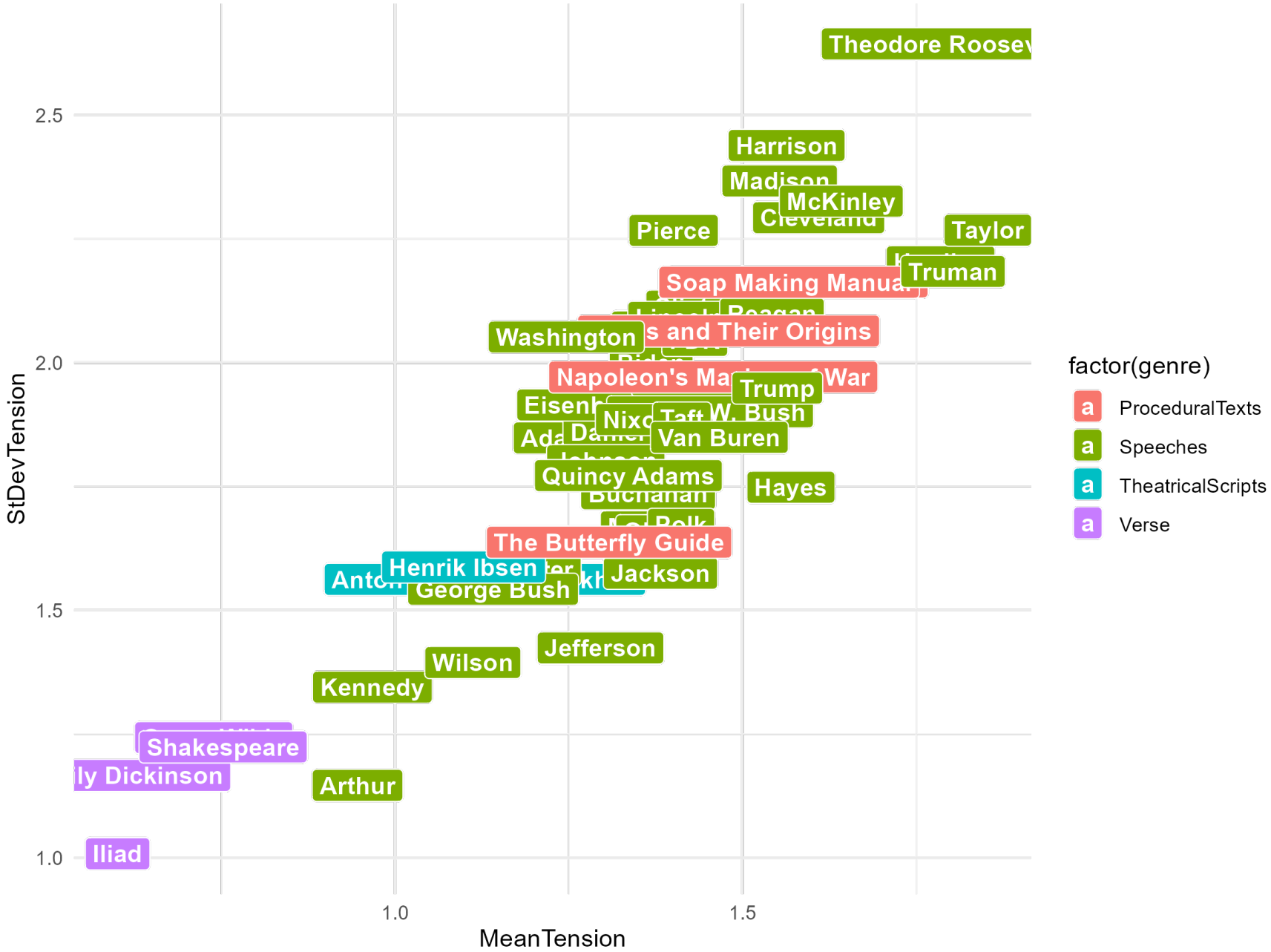
Prose lines with varying metricality

	LINE FRAGMENT	MU	MTS
(a)	infusorial earth of various kinds, silex, etc. (*Soap Making Manual*)	65	502
	…	…	…
(b)	And this may go on indefinitely. (*Criminal Psychology*)	6	26
	…	…	…
(c)	The capillary rise or spread of water (*Rocks and Their Origins*)	1	0

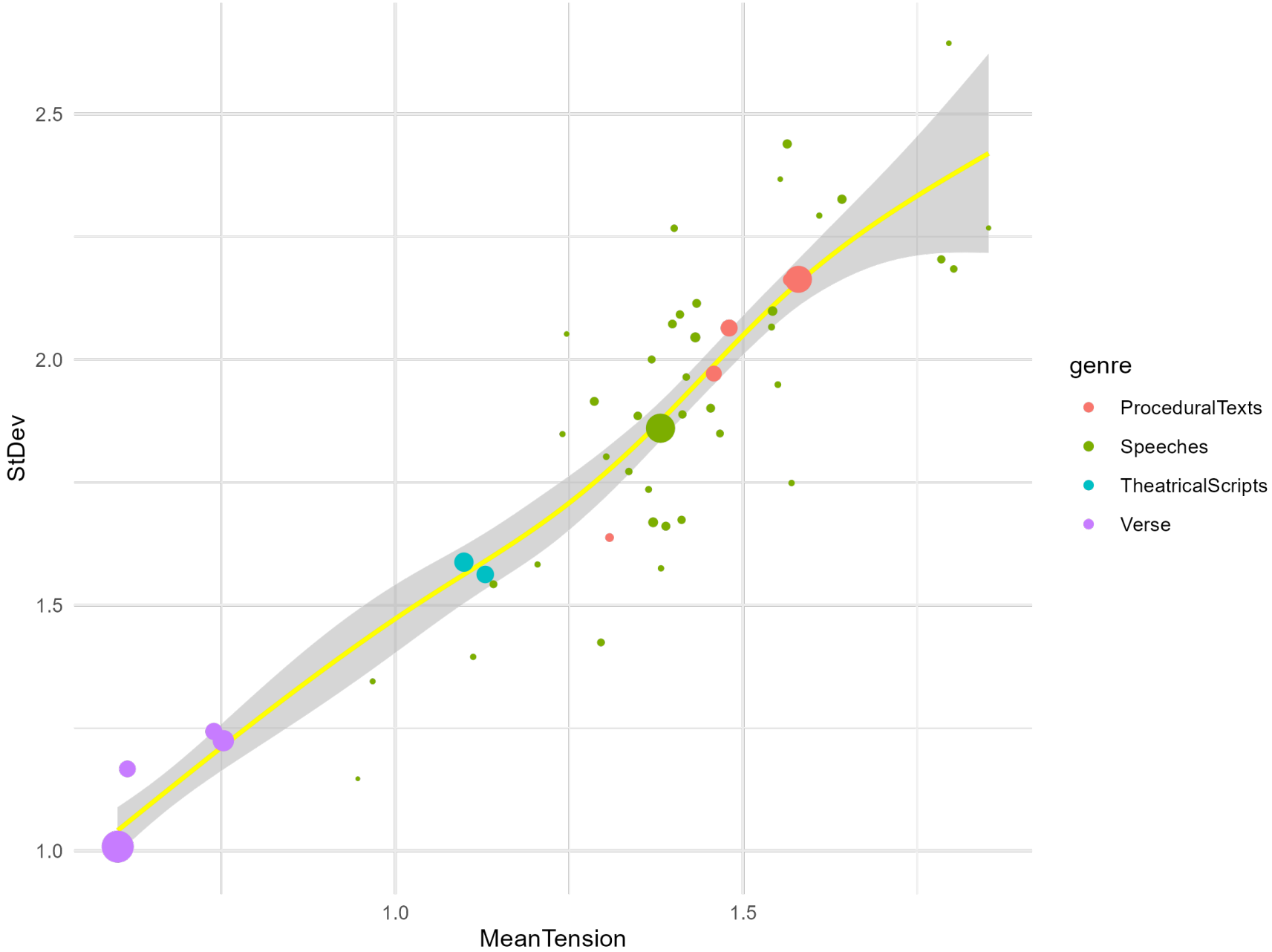
Interpreting Metrical Tension Sum (MTS)

- MTS is the metrical tension sum of an **individual line**.
- Mean MTS is the average metrical tension sum **across lines**.
  - low in Verse
  - higher in Theatrical Scripts and Speeches
  - highest in Procedural Texts
- Standard Deviation of MTS shows how much metrical tension sum varies **across lines**.

51 texts, x = mean(MTS), y = sd(MTS)



51 texts, x = mean(MTS), y = sd(MTS)



Testing the Metrical Closure Hypothesis

- **Metrical Closure**: In metrical verse, beginnings of lines are lax, endings are strict.
- **Our hypothesis**: The same happens in any text.
- Here's an initial linear regression model:

  `lm(tension ~ initial)`, where

`tension` (= MTS)	0 . . . 35.79	(numeric)
`initial`	TRUE or FALSE	(logical)

- **Prediction**: We expect line-initial strings to have higher tension, i.e., `initial` should have a positive coefficient.

Testing the Metrical Closure Hypothesis

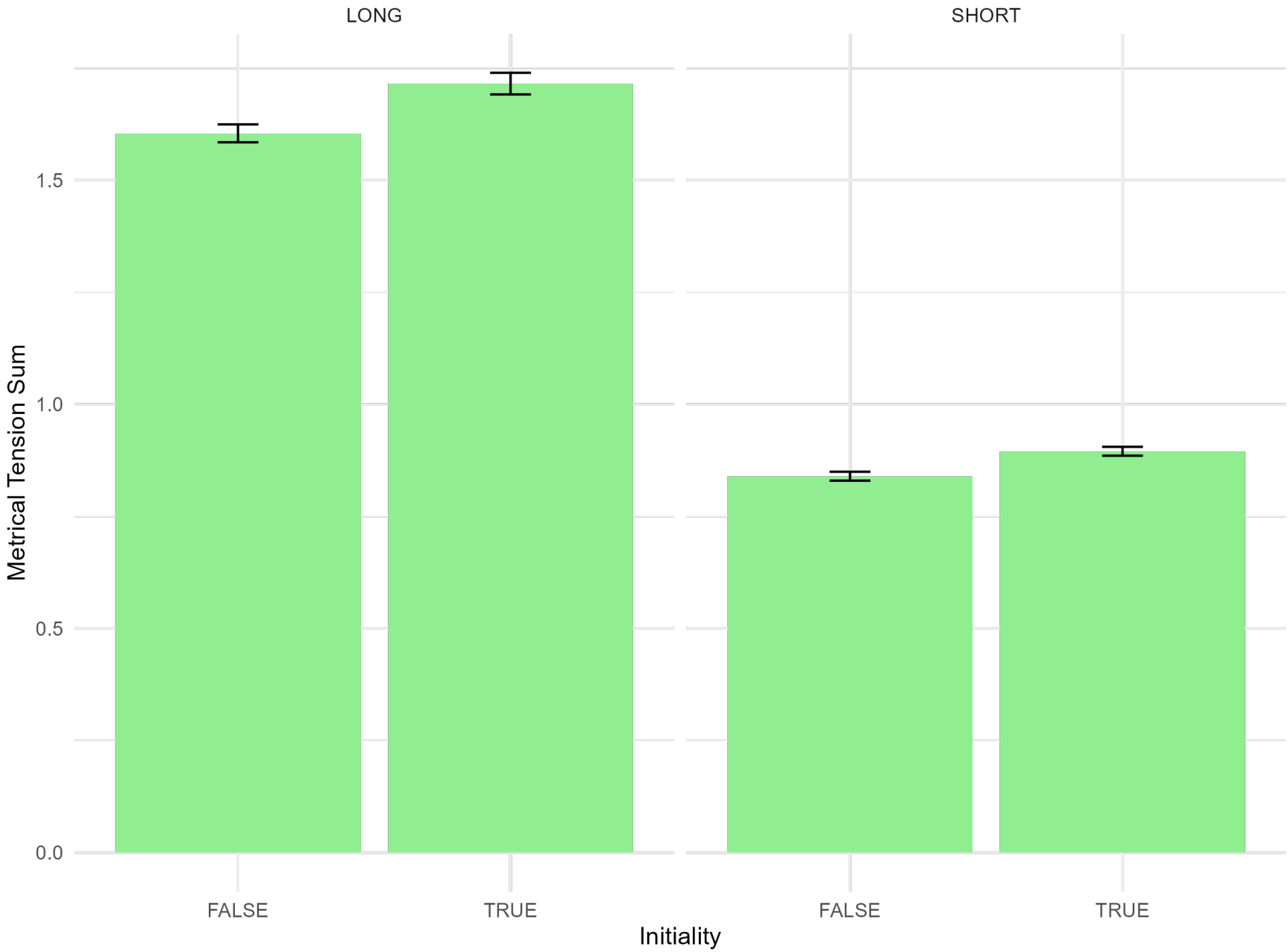
However, this ignores an important variable: word length.

- Prose favors longer words than verse (Anttila and Heuser 2016). This may be for several reasons:
  - differences in topic
  - long words easily violate metrical constraints
- Genres differ in how constituent length matters to linearization (Saintsbury 1912/1965; Croll 1919; Blumenfeld 2016):
  - in utility prose the tendency is long last ("end weight")
  - in verse and art prose the opposite tendency has been observed
- Revised model: We include mean word length as a control.

`lm(tension ~ initial + mean_word_length)`, where

`mean_word_length`	1...5	(numeric)

A visual confirmation



Regression modeling

- The model formula:

```
lm(tension ~ initial + mean_word_length)
```

- Model summary:

initial	$b = 0.11$, $SE = 0.01$, $p < 0.001$
mean_word_length	$b = 0.98$, $SE = 0.02$, $p < 0.001$

- Our hypothesis is supported: MTS is positively associated with both line-initiality and mean word length.
- In other words, metrical tension is higher in line-initial strings and increases with word length, just as expected.

More sophisticated regression models

- We added the fixed effect of `genre` (4 levels, Helmert-coded). The initiality and word length effects persisted ($p < 0.001$).
- We fitted a mixed model with `source` (51 groups) as a random intercept. Both effects persisted ($p < 0.001$).
- Finally, we fitted Poisson regression and negative binomial regression models to the data (Winter 2020, 218-231), with similar results. A likelihood ratio test of a negative binomial model against a Poisson model revealed a significant difference ($\chi^2(1) = 23884.90, p < 0.0001$).

Regression modeling: Discussion

- **The role of word length**
  Interestingly, metrical closure only shows up if we control for word length. Maybe that is why it has not been observed before?

- **The role of phrasal stress**
  What explains metrical closure? Following Ryan (2017), we speculate that metrical constraints become stricter under increasing phrasal stress (Chomsky and Halle 1968); cf. Hayes (1983) who states a similar generalization in terms of prosodic domains.

Conclusion

- Metrical closure is not just a property of metered verse but a general property of language.

Acknowledgements

- This research was partially funded by the Roberta Bowman Denning Initiative in the Digital Humanities at Stanford University as part of the project *Prose Rhythm and Linguistic Theory*.
- We have benefited from interactions with Scott Borgeson, Joan Bresnan, Nigel Fabb, Joshua Falk, Vivienne Fong, Antón de la Fuente, Kristin Hanson, Bruce Hayes, Ryan Heuser, Brian Hicks, John Ioannidis, Paul Kiparsky, Brennan Nick, and Scott Stevens.
- Visualization and statistical analyses were done in R (R Core Team 2024) and `ggplot2` (Wickham 2016).

Thank you!

References I

Anttila, Arto, and Ryan Heuser. 2016. Phonological and metrical variation across genres. *Proceedings of 2015 Annual Meeting on Phonology* .

Anttila, Arto, Ryan Heuser, and Paul Kiparsky. 2022. Prose rhythm and antimetricality. LSA Symposium on Literary Linguistic Forms, January 8, 2022. Slides available at `https://web.stanford.edu/~anttila/`.

Blumenfeld, Lev. 2015. Meter as faithfulness. *Natural Language & Linguistic Theory* 33:79–125.

Blumenfeld, Lev. 2016. End-weight effects in verse and language. *Studia Metrica et Poetica* 3:7–32.

Breiss, Canaan, and Bruce Hayes. 2020. Phonological markedness effects in sentence formation. *Language* 96:338–370.

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. Cambridge, Mass.: MIT Press.

Croll, Morris W. 1919. The cadence of English oratorical prose. *Studies in Philology* 16:1–55.

Day, David. 2023. Syllables: A fast syllable estimator for Python. Available at `https://github.com/prosegrinder/python-syllables/`.

Fabb, Nigel. 1997. *Linguistics and Literature*. Oxford, UK, and Malden, MA: Blackwell Publishers.

de la Fuente, Antón, Brennan Nick, and Arto Anttila. 2023. Metrical uncertainty. Poster presented at *MorrisHalle@100*, September 10, 2023, MIT, Cambridge, MA. Slides available at `https://web.stanford.edu/~anttila/`.

References II

Halle, Morris, and Samuel Jay Keyser. 1971. *English Stress: Its Form, Its Growth, and Its Role in Verse*. New York–Evanston–London: Harper & Row, Publishers.

Hanson, Kristin, and Paul Kiparsky. 1996. A parametric theory of poetic meter. *Language* 72:287–335.

Hayes, Bruce. 1983. A grid-based theory of English meter. *Linguistic Inquiry* 14:357–393.

Hayes, Bruce, Colin Wilson, and Anne Shisko. 2012. Maxent grammars for the metrics of Shakespeare and Milton. *Language* 691–731.

Heuser, Ryan, Joshua Falk, and Arto Anttila. 2010. Prosodic. Software, Stanford University, available at `https://github.com/quadrismegistus/prosodic` and `http://prosodic.stanford.edu/`.

Kiparsky, Paul. 1977. The rhythmic structure of English verse. *Linguistic Inquiry* 8:189–247.

Kiparsky, Paul. 2020. Metered verse. *Annual Review of Linguistics* 6:25–44.

Prince, Alan. 1989. Metrical forms. In *Rhythm and meter*, Phonetics and phonology, Vol. 1, 45–80. Academic Press, Inc.

Prince, Alan, and Paul Smolensky. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, Massachusetts: Blackwell Publishing.

R Core Team. 2024. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL `https://www.R-project.org/`.

Ryan, Kevin M. 2017. The stress–weight interface in metre. *Phonology* 34:581–613.

References III

Saintsbury, George. 1912/1965. *A History of English Prose Rhythm*. Bloomington: Indiana University Press. First edition 1912.

Smith, Barbara Herrnstein. 1968. *Poetic Closure*. Chicago and London: The University of Chicago Press.

Steele, Timothy. 1999. *All the Fun's in How You Say a Thing: An Explanation of Meter and Versification*. Athens: Ohio University Press.

Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. URL `https://ggplot2.tidyverse.org`.

Winter, Bodo. 2020. *Statistics for linguists: An introduction using R*. Routledge.

Woolley, John T., and Gerhard Peters. 1999–. The American Presidency Project. Santa Barbara, California: UC Santa Barbara.