

# The Rise and Fall of Antimetricality

Scott Borgeson Arto Anttila Ryan Heuser Paul Kiparsky

Stanford University

## Abstract

“Should we not, Monsieur, carefully avoid Alexandrines in prose?” So asks the Self-Taught Man in Sartre’s *Nausea*, pointing to the traditional view that, rhythmically, prose is prose by avoiding meter. Indeed, for Saintsbury in *A History of English Prose Rhythm* (1912), the “great law” of prose is that “every syllable shall, as in poetry, ... be capable of entering into rhythmical transactions with its neighbours, but that these transactions *shall always stop short ... of admitting the recurrent combinations proper to metre*.” This paper traces such rhythmical tensions between prose and verse across English-language literary history. We apply to a corpus of prose and verse a set of new computational tools for metrical scansion, which measure the extent to which the phonological features of written text can be mapped onto a metrical grid. Our goal is to test Saintsbury’s “great law,” along with a sharpened form of it which, drawing on Jakobson, we call the RELATIVIZED ANTI-METRICALITY HYPOTHESIS: namely, that meter is inscribed as a negative presence in the literary prose of a given period to the extent that metrical verse is then the dominant literary form. On such a view, prose actively avoids metricality during the dominance of verse between the sixteenth century and the nineteenth (e.g. Browne, Addison); during this period, then, prose can be called “anti-metrical.” Then, in the nineteenth century, as the dominance of verse is eclipsed by the rise of the novel, literary prose starts to flirt with meter (e.g. Dickens, Ruskin), thus explicitly opposing its former mandate and becoming, instead, “anti-anti-metrical.” Finally, as metrical verse collapses in the twentieth century, and metricality as a rhythmic posture fades from literary view, prose abandons all relationship to meter, whether positive or negative, to become instead “a-metrical.”

We present here preliminary evidence in support of this hypothesis about rhythm’s literary history. We examine five literary authors from different periods: Shakespeare, Browne, Dickens, Ruskin, and Pater, along with Samuel Dibble, author of *Elements of Plumbing* (1918), our baseline text. We have three main results. First, we find a metricality continuum which broadly agrees with Saintsbury’s and our own judgments. Second, based on experiments with various metricality measures, we propose the ACCIDENTAL PENTAMETER INDEX, METRICAL FRICTION, and METRICAL UNCERTAINTY as promising new metrics for measuring prose rhythm. Finally, we find that when these metrics are applied to thousands of prose and verse texts across several centuries of literary history, the results are consistent with the relativized antimetricality hypothesis.

## 1. Introduction

Two principles of Jakobsonian poetics, MARKEDNESS and DOMINANCE, lead jointly to global predictions that are hard or impossible to verify by traditional methods.<sup>1</sup> Newly developed

---

<sup>1</sup> This work started as a collaborative effort during the course LINGUIST 213 Corpus Phonology: Prose Rhythm at Stanford University in the Spring Quarter of 2017. We thank Joan Bresnan, Vivienne Fong, and Scott Stevens for their input, and the Department of Statistics Consulting Services at Stanford University for statistical advice. The

automatic scansion tools that detect metrical properties of texts on a large scale have now made it possible to do so. This essay is a first report on our ongoing project to explore English prose rhythm, and in particular the phenomenon of ANTIMETRICALITY in literary prose and its variation across authors and periods. At the same time, our study also serves as a test of the ability of these as yet untried tools to deliver consistent, accurate, and revealing results.

The principle of markedness is that formal categories are structured by a hierarchy of asymmetrical oppositions.<sup>2</sup> It conceives of the forms of verbal art, like those of language itself, as arising by a series of successive differentiations, each resulting in an increment of complexity within the superordinate category. The first marks off literary language from ordinary language by imposing on it regular patterns of recurrence, such as meter and parallelism. These define verse as the intrinsically unmarked form of literary language. Literary prose is then marked in turn by the suppression of some or all such verse features, which puts it conceptually at two removes from ordinary language: it represents the marked category of non-ordinary language, and within that category the marked subcategory of non-verse. Thus, verse is already inscribed as a negative presence in all literary prose.

The idea of prose as anti-meter has respectable pretheoretical antecedents. Saintsbury's extremely perceptive *History of English Prose Rhythm* (1912) states that the foot-scansion of prose "is arranged on a principle totally different, and indeed opposed, when compared with that of poetry. Instead of sameness, equivalence, and recurrence, the central idea turns on difference, inequality, and variety. And though a certain amount of correspondence is introduced by the necessary presence of the identical quantity-combinations called feet, *these are to be so arranged that they will not constitute metre*" (p. 344). The "great law" of prose is "that every syllable shall, as in poetry, have recognisable rhythmical value, and be capable of entering into rhythmical transactions with its neighbours, but that *these transactions shall always stop short, or steer clear, of admitting the recurrent combinations proper to metre.*" "Thus rhythmical prose, in its perfection, is distinguished from poetry by subtle but easily recognisable differences of diction, arrangement, and the like, but most of all, and most essentially, by *the absence of definite and ostentatious correspondence in rhythmical-metrical character, and of equivalent or definitely corresponding 'lines'.*"

As a normative principle of literary prose, antimetricality traces back to classical rhetoric. According to Quintilian (9.4.72) "an entire verse in prose is the most disgusting (*foedissimum*), but even a part is ugly, especially in the final cadence or at the beginning." Exploratory work by Kevin Ryan (p.c.) suggests that the practice of Latin authors is consistent with his judgment. He

---

development of the Prosodic software was financially supported by the Hellman Junior Faculty Scholar Fund, the Office of the Vice Provost for Undergraduate Education (VPUE), and the Roberta Bowman Denning Initiative Committee, H&S Dean's Office, as part of the project *Prose Rhythm and Linguistic Theory*. We are responsible for any errors.

<sup>2</sup> Jakobson 1935a, 1941. On dominance, see Jakobson 1935b. The connection between markedness and dominance and its implications for the theory of prose and mixed forms is explored in Hanson and Kiparsky 1997, on which we draw here.

found 137 accidental hexameters in 55,751 words of Latin prose. A para-text he constructed by randomizing the order of the words contained 199 hexameters chunks, almost 50% more in the same number of words, a statistically significant difference.<sup>3</sup>

The French rhetorician Vaugelas (1647) expresses the same injunction: “one must avoid verses in prose as far as possible, especially alexandrines and common verses.” French prose writers have until recent times tried to abide by this norm (Purnelle 2005).

Is Saintsbury right that English literary prose is subject to a similar constraint? His intuition is entirely plausible, but we are not aware of any systematic attempt to verify it. Saintsbury asserts that literary prose contains fewer accidental blank verses than expected by chance, and cites cases where English prose writers seem to choose their syntax and vocabulary so as to avoid lapsing into meter. But impressionistic generalizations and hand-picked examples cannot prove the point. Systematic statistical comparison of literary prose with utility prose on a large scale is required. This is much harder to do in English than in Latin or French, because English writing is not a reliable guide to stress, which is the basis of English meter. It has only now become possible with automatic metrical annotation and scansion technology such as we use in our study.

Our study also aims to test a further prediction, which is derived by combining the principle of markedness with the principle of DOMINANCE. Markedness entails that the terms of a binary opposition are never simply interchangeable. Verse and prose cannot exist side by side in a stylistic equilibrium. Default expectations about them are exploitable for compositional ends, in a specific historical context partially determined by those hierarchical and asymmetrical relations themselves. These expectations include the dominance of one form, normally verse at the beginning of a literary tradition and prose at the end, and the primacy of verse for lyric functions and of prose for narrative.

While markedness is an intrinsic invariant property of a specific term in a formal opposition, dominance is a contingent shifting valorization of one of its terms, which defines a literary period. Being the unmarked form of literary language, verse is the dominant form of unwritten and early written literatures in all genres, whether in its unmarked function of lyric, or in the marked function of epos and drama. The majority of oral literatures to this day are verse.<sup>4</sup> New vernacular literatures of emerging nations mark their modernity by a shift in dominance from verse to prose. This shift took place in Western Europe beginning in the 13th century, in Eastern Europe in the 19th century, and is an ongoing process in many new post-colonial national literatures today. In English literature, the new dominance of prose lasted from the decline of

---

<sup>3</sup> The search included all possible subsequences of words. Real hexameters have to coincide with phrase boundaries. Ryan notes that imposing that requirement would yield many fewer accidental hexameters and probably reveal a stronger avoidance effect.

<sup>4</sup> The vehicle of Indo-European literature was verse in iambic octosyllables combinable into distichs and quatrains; even the hexameter turns out to be an offspring of this form (Kiparsky 2018). Uralic literature was likewise based on verse, with parallelism as its major constitutive device; trochaic octosyllabic parallel distichs can be reconstructed as the form in one branch of the family.

alliterative verse through the Middle English and early Modern English period into the early 16th century, when another wave of formal renewal returned verse to dominance, with the freshly reactivated iambic meter as its vehicle. It begins to crumble again in the late 18th century, declining further as the novel rises in importance, and in spite of innovations (Sprung Rhythm, free verse, hip hop) remains a secondary form.

Shifts between the dominance of verse and prose are reflected in literary practice and theory. Translation is normally into the dominant; so is the paraphrasing of older works within a language.<sup>5</sup> A new dominant is also overtly legitimized by claiming for it either the virtues that were earlier attributed to the opposite one, or superior virtues of its own.<sup>6</sup> The architecture of mixed forms (prosimetrum) reveals the dominant in a particularly transparent way (Hanson and Kiparsky 1997).

Putting together the ideas of markedness and dominance, we arrive at the RELATIVIZED ANTI-METRICALITY HYPOTHESIS. This is that the extent of antimetricality effects in literary prose are prevalent when verse is dominant, and attenuated when it is recessive. It predicts that antimetricality effects predominate in medieval vernacular writing ca. 900-1300, and then again from the 16th through the 18th centuries, and fade from the 19th century to the present.

A perusal of Saintsbury's intuitions provides suggestive *prima facie* support for the relativized antimetricality hypothesis. His prime examples of antimetrical prose writers are Donne ("absolute perfection of rhythmical — never metrical — movement", p. 163), Browne ("variety pushed to what might have seemed antecedently an almost impossible point", p. 191), Addison ("mastered the great principle of variety", p. 248), and Pater ("far more intricate and nuanced than in Ruskin or Kingsley", p. 425), who according to him have hardly any accidental blank verses in their prose. All except Pater fall squarely in the period when verse is dominant.

---

<sup>5</sup> Up to the 13th century, Latin prose was normally turned into French vernacular verse (Godzich and Kittay 1987). With the 13th century reversal of dominance, prose translations of Latin become popular, even from verse and mixed forms, as in Jean de Meun's translation of Boethius. The *dérimage* (de-versification) of older French verse into vernacular prose flourished at the same (Spiegel 1993). These practices continued up to the 16th century. From then on, prose translations of verse become rare for several centuries. (Anne Dacier's prose French Homer (1699, 1708) is a notable exception.) In the 19th century, verse is routinely translated into French prose again.

The same trends can be seen in English. From the 16th to the 18th century, all translation of Homer into English were in verse, including the famous ones by Chapman, Ogilby, Hobbes, Dryden, and Pope. (MacPherson's 1773 translation, though not lineated, is a kind of Ossianic crypto-verse.) The 19th century saw no less than ten English prose translations of the Iliad alone, and their production has continued unabated since then (see [johnstoniatexts.x10host.com/homer/homertranslations.htm](http://johnstoniatexts.x10host.com/homer/homertranslations.htm); [en.wikipedia.org/wiki/English\\_translations\\_of\\_Homer](http://en.wikipedia.org/wiki/English_translations_of_Homer)).

<sup>6</sup> The Greek rhetoricians claim that prose has a style which excels that of verse in ornamentation and figurative richness, and a rhythm which excels that of verse in variety. French apologists for prose consider it more dignified than verse, and hence more appropriate for certain historical themes (Godzich and Kittay, 150). Dante in *Il Convivio* (I.10.12) argues that Italian prose is capable of expressing concepts "which could not be expressed appropriately in rhyme, because of the accidental adornments which are connected with it, namely rhyme and regulated number, in the same way as the beauty of a woman cannot be seen when the adornments of fancy dress are the cause of her being admired more than she herself is. In Johannes von Tepl's *Ackermann aus Böhmen*, (c. 1401) Death acknowledges the plowman's choice of prose as a sign of his serious intent.

- (1) That these were the urns of Romans from the common custom and place where they were found, is no obscure conjecture, not far from a Roman garrison, and but five miles from Brancaster, set down by ancient record under the name of Branodunum. And where the adjoining town, containing seven parishes, in no very different sound, but Saxon termination, still retains the name of Burnham, which being an early station, it is not improbable the neighbour parts were filled with habitations, either of Romans themselves, or Britons Romanized, which observed the Roman customs. (Thomas Browne, *Urn-Burial* [1658])

Pater's self-consciously classicizing style harks back to an earlier period, and could reasonably be considered a special case.

- (2) To reconcile forms of sentiment which at first sight seem incompatible, to adjust the various products of the human mind to each other in one many-sided type of intellectual culture, to give humanity, for heart and imagination to feed upon, as much as it could possibly receive, belonged to the generous instincts of that age. An earlier and simpler generation had seen in the gods of Greece so many malignant spirits, the defeated but still living centres of the religion of darkness, struggling, not always in vain, against the kingdom of light. Little by little, as the natural charm of pagan story reasserted itself over minds emerging out of barbarism, the religious significance which had once belonged to it was lost sight of, and it came to be regarded as the subject of a purely artistic or poetical treatment. (Walter Pater, *The Renaissance: Studies in Art and Poetry* [1873])

Saintsbury's prime examples of "anti-anti-metrical" writers who on the contrary flirt with meter are from periods where prose is dominant: Dickens ("apt to abuse [blank verse]", Saintsbury 1912: 381), and Chaucer, who indulges in iamboid prose with lines of irregular length ("a mistaken attempt to adorn prose by calling in the direct aid of metre", 1912: 71)—and no one in the centuries between them.

- (3) When the day broke, it blew harder and harder.  
I had been | in Yar|mouth when | the sea|men said  
it blew | great guns, | but I | had ne|ver known  
the like | of this, | or a|nything | approach|ing  
to it. | We came | to Ips|wich—ve|ry late,  
having | had to | fight ev|ery inch | of ground  
since we | were ten | miles out | of London; (Dickens, *David Copperfield* [1850])
- (4) A young | man call|ed Me|libeus, migh|ty and rich, | begat  
Upon | his wife, | that call|ed was | Prudence,  
A daugh|ter which | that call|ed was | Sophie.  
Upon | a day | befell  
That he | for his | disport | is went into

The field | es him | to play.  
 His wife | and eke | his daugh|ter hath | he left  
 In with | his house, | of which | the do|res we|ren fast | yshette.  
 Three of | his ol|de foes |it han | espyed . . . (Chaucer, “Tale of Melibee” [ca. 1373])

The case of Ruskin is somewhat different. “You can hardly open a page of Ruskin—when his prose has caught fire—without finding [blank verses]” (Saintsbury p. 395):

(5) A mul|titude | of pil|lars and | white domes,  
 clustered | into | a long | low py|ramid  
 of col|oured light; | a trea|sure-heap, | it seems,  
 partly | of gold, | and partly | of o|pal and  
 mother-|of-pearl, | hollowed | beneath | into five  
 great vaul|ted por|ches, ceiled | with fair | mosaic,  
 and beset | with sculpture | of a|llabas|ter, clear  
 as amber | and del|icate | as i|vory, (Ruskin, *Stones of Venice* [1851])

But most of the time it has not caught fire, and looks like this:

(6) Now there is apparently a significative coincidence between the establishment of the aristocratic and oligarchical powers, and the diminution of the prosperity of the state. But this is the very question at issue; and it appears to me quite undetermined by any historian, or determined by each in accordance with his own prejudices. It is a triple question: first, whether the oligarchy established by the efforts of individual ambition was the cause, in its subsequent operation, of the Fall of Venice; or (secondly) whether the establishment of the oligarchy itself be not the sign and evidence, rather than the cause, of national enervation; or (lastly) whether, as I rather think, the history of Venice might not be written almost without reference to the construction of her senate or the prerogatives of her Doge. (Ruskin, *Stones of Venice* [1851])

The purple patches like (5) that Saintsbury notices are a hallmark of anti-antimetricality. But our Prosodic does not single them out. Based on the preponderance of prose like (6) in the Ruskin’s text, it resolutely puts him among the antimetrical prose writers.

The incidence of antimetricality, then, appears to be at least in rough agreement with the relativized antimetricality hypothesis. However, new automatic scansion technology can help us investigate it more broadly and empirically.

A corollary of relativized antimetricality is that both antimetrical and metrical effects should entirely disappear in prose that aspires to eliminate all artifice, such as emerges in certain modernist styles. Under the banner of that ideology, literary prose should become AMETRICAL, formally merging with utility prose. A writer who introduced a new plain prose style into French prose is Sartre. His *Nausea* contains the following revealing passage.

- (7) A great, gaunt shadow suddenly springs up behind me. I jump.  
 “Excuse me, Monsieur, I didn’t mean to disturb you. I saw your lips moving. You were undoubtedly repeating passages from your book.” He laughs. “You were hunting Alexandrines.”  
 I look at the Self-Taught Man with stupor. But he seems surprised at my surprise:  
 “Should we not, Monsieur, carefully avoid Alexandrines in prose?”<sup>7</sup>

Sartre here takes care to announce that he knows the classical norm, and has decided to simply ignore it. In fact, *Nausea* neither obeys the rule that forbids alexandrines, nor ostentatiously flouts it by an excess of them. Accidental alexandrines occur in the text with the same random frequency as in utility prose (Purnelle 2005).

In order to test the antimetricality hypothesis, both in its general form and in its relativized form, we have to make it more precise. On a literal understanding of Quintilian’s and Vaugelas’ formulations, and of some of Saintsbury’s statements too, antimetricality can be satisfied by just avoiding accidental prototypical metrical lines in a text: hexameters in Latin, alexandrines in French, blank verse in English. Such a requirement — call it WEAK ANTIMETRICALITY — can be met by the Self-taught Man’s method ridiculed by Sartre. All that is needed is to disrupt an otherwise metrical sequence by a single metrical violation, such as a misplaced word stress or an excess syllable. This is a rather uninteresting version of antimetricality, but it makes some sense if we think of antimetricality as an artifice consciously used by literary prose writers. To detect Weak Antimetricality, the subject of the experiment in section 3.1, we developed automatic scansion tools to count the number of times a stretch of blank verse *runs up against a punctuation mark*, like a comma or semicolon, closing a phrase.

A subtler version of antimetricality is hinted at in Saintsbury’s mention of the “absence of definite and ostentatious correspondence in rhythmical-metrical character, and of equivalent or definitely corresponding ‘lines.’” On at least one reading, this means that good prose writers do not just avoid accidental blank verses, but actually seek out rhythmic complexity and interest in their prose. This makes antimetricality independent of any particular meter. It is violated as much by an accidental anapestic trimeter as by blank verse, or by any other regular rhythm for that matter. Moreover, it becomes a *gradient* property which can be satisfied to a greater or lesser extent in a given text or language. The avoidance of blank verse would then be just a side effect of rhythmic variation and complexity. We have therefore also developed automatic scansion tools to detect such STRONG ANTIMETRICALITY in a text. These tools impose thousands of possible metrical scansions onto a string of text, and then return only the possibilities that can’t be ruled

---

<sup>7</sup> Une grande ombre maigre surgit brusquement derrière moi. Je sursaute.

– Excusez-moi, monsieur, je ne voulais pas vous déranger. J’ai vu que vos lèvres remuaient. Vous répétiez sans doute des phrases de votre livre. – Il rit. – Vous faisiez la chasse aux alexandrins. Je regarde l’Autodidacte avec stupeur. Mais il a l’air surpris de ma surprise:

– Ne doit-on pas, monsieur, éviter soigneusement les alexandrins dans la prose?

out according to a few simple rules. As we will see, the number of such remaining, possible scansions — what we call METRICAL UNCERTAINTY — effectively measures Strong Antimetricality.<sup>8</sup>

## 2. Method

The question of how to distinguish a metrical line from an unmetrical line is at the heart of the study of metrics.<sup>9</sup> Closely related and no less interesting is the question of how to distinguish a line of metered verse from a line of unmetered prose. To get a sense of the problem, consider the following random five-word text fragments organized into two columns. One column is verse, the other column is prose. The task is to determine which is which (we invite you to guess before continuing onto the next paragraph).

(8) mankind do know of hell	readiness to measure time by
fled away into the storm	in a trio while i
her vespers done of all	the weather is unfavourable for
a richness that the cloudy	be in time perhaps it
fix'd as in poetic sleep	i shall horribly commit myself
cold fair isabel poor simple	as bad again just now
little cottage i have found	i shall have got some
last prayer if one of	bless you sunday evening my
one hour half-idiot he stands	bars at charles the first

Both columns were extracted from the works of John Keats. The fragments on the left are verse; the fragments on the right are prose, taken from his letters. Distinguishing prose from verse no doubt involves many factors, including typography, vocabulary, parallelism, and rhyme. In the present study, we set ourselves the goal of understanding to what extent basic phonological notions can help classify a fragment of text as metered verse or unmetered prose. In some cases the role of phonology is fairly obvious. To continue with one of Steele's examples, it is not hard to hear that the ten-syllable sentence *I can't believe that I forgot my keys* works well as a line of iambic pentameter, whereas another ten-syllable sentence *It rains almost always when I visit* does not. This judgment likely has something to do with the distribution of word stresses in these sentences, as can be seen from the scansions in (2).

---

<sup>8</sup> Let us add that analogous questions arise for free verse. Does the ghost of some simple meter lurk behind its arras, as Eliot claimed? We now have the tools at hand for answering that question as well; we do not yet have the data, however. See below, section 3.3, for more information.

<sup>9</sup> This section is based upon Anttila and Heuser 2016, which can be consulted for more details.



(8)		(w s)	(w s)	(w s)	(w s)	(w s)		(w s)	(w s)	(w s)	(w s)	(w s)
	(a)	I cán't believe that I forgót my kéys						(b)	It ráins álmóst álwáys wén I vísit			
			s		w				s		w	
		+ stress	4		0			+ stress	1		4	
		– stress	1		5			– stress	4		1	

The Steele quote suggests that in stress-based meters stressed syllables tend to avoid weak metrical positions and unstressed syllables tend to avoid strong metrical positions. This means that we may be able to use the degree of alignment between strong vs. weak positions and stressed vs. unstressed syllables as a crude first approximation of the metricality of a scansion. In (8a) all stressed syllables are neatly aligned with strong (even-numbered) positions and all unstressed syllables are almost as neatly aligned with weak (odd-numbered) positions except for the lexically unstressed pronoun *I* which fills a strong position. In the second example the alignment pattern is nearly reversed. This accounts for the judgment that (8a) works well as a line of iambic pentameter whereas (8b) does not. The view implicit in this illustrative example is that metricality has to do with the correspondence between metrical positions (s, w) and their phonological realizations.<sup>10</sup>

## 2.1 Metrical constraints

In this paper, we make use of a computational tool that automates this process of analyzing the metrical difference between lines such as (8a) and (8b) above. This tool assesses the extent to which the phonological features of language (mainly stress) fit neatly into a metrical grid of weak and strong positions. To do so, it relies on a large body of work in generative metrics that theorizes meter as a mapping of metrical templates to linguistic patterns of prominence. Regulating the relationship between meter and stress (English's primary linguistic pattern of prominence) are a set of correspondence "constraints," effectively rules or preferences that state, for instance, that unstressed syllables should not appear in strong metrical positions, or that stressed syllables should not appear in weak metrical positions. Prosodic works by imposing a wide variety of possible metrical templates on the phonology or stress pattern of a line, thereby assessing which of these templates violates the fewest such correspondence constraints. By these means, the best "fit" metrical template is determined. In addition, the degree of fit is also recorded by the total number of constraint violations even the best fit metrical template encountered.

Which constraints did we use in our analysis? We adopt five constraints from the work of Hanson and Kiparsky (1996), henceforth H&K. These constraints are listed below in (9).

<sup>10</sup> See, e.g., Kiparsky 1977; Prince 1989; Hanson and Kiparsky 1996; Hayes, Wilson and Shisko 2012; and Blumenfeld 2015. Golston 1998 and Golston and Riad 2000, 2005 present an alternative view.

- (9) a. \*W/PEAK (“The Shakespeare Constraint”)  
‘A weak position should not contain strong syllables (“peaks”).’  
b. \*S/UNSTRESSED (“The Hopkins Constraint”)  
‘A strong position should not contain unstressed syllables.’  
c. \*W/STRESSED  
‘A weak position should not contain stressed syllables.’  
d. W-RESOLUTION  
‘For disyllabic positions within a word, the first position must be light and stressed.’  
e. F-RESOLUTION  
‘A disyllabic position crossing a word boundary must be weak and its two words must be function words.’

A PEAK is defined as the head of a branching constituent, roughly, the main stress of a polysyllable; a TROUGH is defined as a non-head of the same. Thus, the words *mány* and *réptile* have the structure *peak* + *trough* and the words *imménse* and *màintáin* have the structure *trough* + *peak*. The monosyllabic word *kéen* is neither a peak nor a trough. To illustrate the constraint \*W/PEAK consider the examples in (5). In canonical English verse violations of this constraint are routine in the first foot (inversion) and sometimes occur after major syntactic breaks within the line, but are avoided elsewhere.

(11)		Number of *W/PEAK violations
	<div style="display: flex; justify-content: space-between; margin-bottom: 5px;"> <span>(w s) (w s) (w s) (w s) (w s)</span> </div> <div style="display: flex; justify-content: space-between;"> <span>Néver</span> <span>cáme</span> <span>póison</span> <span>fróm</span> <span>só swéet</span> <span>a place</span> </div>	1 <i>(Richard III.1.2)</i>
	<div style="display: flex; justify-content: space-between; margin-bottom: 5px;"> <span>(w s) (w s) (w s) (w s) (w s)</span> </div> <div style="display: flex; justify-content: space-between;"> <span>Néver</span> <span>had rát-</span> <span>póison</span> <span>só swéet</span> <span>a taste</span> </div>	2 (construct)

According to H&K, mainstream English and Finnish meters differ crucially with respect to these two constraints (H&K, pp. 287-8): Shakespeare’s iambic pentameter, and the dominant tradition of English metrical verse, avoids violations of \*W/PEAK; Finnish iambic-anapestic (trochaic-dactylic) meters, as well as Hopkins’ sprung-rhythm, avoid violations of \*S/UNSTRESSED.

By applying these constraints to both verse and prose, we aim to test our hypothesis about the changing metrical distance between verse and prose across different historical periods. We argue that the extent to which prose violates these well-known constraints regulating metrical verse measures the extent to which the rhythms of prose conform to the rhythms of verse.

## 2.2 Automatic scansion

With the recent availability of metrically annotated texts and computational tools for detecting meter, the study of prose rhythm is ripe for a revival. Instead of hand-annotation (see, e.g., Hayes, Wilson, and Shisko 2012) we used PROSODIC (Heuser, Falk, and Anttila 2017), a software package for the phonological analysis and metrical scansion of texts.<sup>11</sup> The obvious advantage of machine analysis is that it is cheaper and faster than hand-annotation and opens up much larger datasets. In our study, the crucial advantage is that Prosodic is able to scan prose as well as verse. It simply does not know the difference. While less accurate than hand-coding, machine scansion is irreplaceable in allowing one to scan both prose and verse using the same fully explicit principles. The machine-annotated data yield rich information about metrical differences among texts as language, regardless of their genre or other modes of categorization.

The input to Prosodic consists of constraints parametrized by the user and plain text entered from the keyboard or from a text file. The output is the same text enriched by phonological and metrical information, including word stress, syllable structure, possible metrical scansions, and for each scansion, the violation count for each constraint. The phonological analysis of English builds on the CMU Dictionary (Weide 1998) and ESPEAK ([espeak.sourceforge.net](http://espeak.sourceforge.net)). Sample annotations are shown in (12). The first column shows the word as it occurs in the text; the second column is the phonological transcription; the third column shows the word stress profile (P = primary stress, S = secondary stress, U = unstressed); and the fourth column shows the syllable weight profile (H = heavy, L = light).

(12)	I	P:'aɪ	S:P	W:H
	shall	P:'ʃæl	S:P	W:H
	horribly	P:'hɔ:.rə.bli:	S:PUU	W:HLH
	commit	P:kə.'mɪt	S:UP	W:LH
	myself	P:ˈmaɪ.'self	S:SP	W:HH

Prosodic does metrical scansion by starting from all possible scansions for a line of input. Since each syllable has the possibility of acting as a weak (w) or a strong (s) position in a metrical template, the upper bound of possibility for a ten syllable line is  $2^{10} = 1,024$  scansions. For instance, taking the stress assignments from (12), Prosodic will attempt to fit the three metrical templates (13a-c), as well as the other 509 ( $2^9 - 3$ ) other possibilities for this 9-syllable line.

(13)	Stress:	PPPUUUPSP
	a. Meter?:	WSWSWSWSW
	b. Meter?:	WWSWSWWS

<sup>11</sup> For more information about Prosodic, and to try it out for yourself, visit its (beta) website at <http://prosodic.stanford.edu>.

c. Meter?: SWSWSWSWS

For all of these possible metrical templates, the correspondence constraints are applied. Does, for instance, a stressed syllable fall in a weak position, thus violating the correspondence constraint \*W/STRESSED? Consider the familiar line *To be or not to be that is the question*. If we require that each metrical position must be occupied by a single syllable, Prosodic only delivers the iambic scansion.

(14)	1	w	to	
	2	s	BE	*S/UNSTRESSED
	3	w	or	
	4	s	NOT	
	5	w	to	
	6	s	BE	*S/UNSTRESSED
	7	w	that	
	8	s	IS	*S/UNSTRESSED
	9	w	the	
	10	s	QUE	
	11	w	stion	

Native speakers sometimes experience individual lines as fitting more than one meter. In addition to the iambic pentameter scansion this line also easily scans as dactylic tetrameter (Blumenfeld 2015, 84). If we relax the meter by allowing a weak position to contain up to two syllables (i.e., resolution, Kiparsky 1989) we get the dactylic scansion as well.<sup>12</sup>

(15)	1	s	TO	*S/UNSTRESSED
	2	w	be or	
	3	s	NOT	
	4	w	to be	
	5	s	THAT	
	6	w	is the	
	7	s	QUE	
	8	w	stion	

What happens now if we turn from examples of verse to prose? The key advantage of Prosodic is that it blindly analyses any text, metered verse as well as unmetered prose. This is crucial if we want to compare verse and prose. Perhaps we can expect verse to have fewer constraint

---

<sup>12</sup> We also get a third scansion with a mixture of iambic and dactylic feet.

violations than prose across the board. After all, poetry is sometimes defined as the best words in the best order. This turns out to be true for some constraints, but not for others. It would also not be surprising to find that prose occasionally scans like metered verse. This can be illustrated by the following sentence from Franklin D. Roosevelt’s first inaugural address which scans as iambs with no violations:

(16)	1	w	the
	2	s	ONL
	3	w	y
	4	s	THING
	5	w	we
	6	s	HAVE
	7	w	to
	8	s	FEAR
	9	w	is
	10	s	FEAR
	11	w	its
	12	s	ELF

The effect is probably intentional.<sup>13</sup> Note that the syntactic variant *Fear itself is the only thing we have to fear* (construct) is metrically worse: parsed iambically, *Fear it / self is /* incurs six constraint violations: \*W/STRESSED (2 violations, *fear*, *-self*), \*S/UNSTRESSED (2 violations, *it*-, *is*), and \*W/PEAK (1 violation, *-self*).

### 2.3 Data

In our primary experiment, we studied Shakespeare’s *Sonnets* and five prose texts: Browne’s *Religio Medici* (1643) and Pater’s *Two Early French Stories* (1872), claimed to contain significantly fewer accidental blank verses than expected by chance; Dickens’ *David Copperfield* (1850) and Ruskin’s *Stones of Venice* (1851), claimed to contain significantly more of them than expected by chance; and finally Dibble’s *Elements of Plumbing* (1918), a text representing utility prose by an author with no metrical agenda.

Following the method in Anttila and Heuser 2016, we converted the texts into lineated versions where each line consists of five consecutive words. Specifically, we divided the text into its sentences, and then divided each sentence into 5-word units, starting from the end of the sentence and moving backward, and discarding any remainder. This is to guarantee that any metrical differences that might emerge among the texts would have nothing to do with sentence

---

<sup>13</sup> For more examples, see Allen 2009. We are indebted to Stephanie Shaw for this reference.

length, but only the choice and order of words. This resulted in a dataset of approximately 3,000 five-word pseudo-lines from each author. We also produced two pseudo-texts: a RANDOMIZED text that replaces words with words of the same part of speech, keeping the syntax relatively intact, and a SCRAMBLED text that disregards part of speech, destroying the syntax. For an example, see Table 1 below.

All in all, our dataset contains 54,949 pseudo-lines. We have 6 authors and about 9,000 lines per author. These lines consist of three text types: 3,000 original, 3,000 randomized, and 3,000 scrambled. Each line has 5 words. The number of syllables per line varies from 5 to 25. Each line was annotated for 35 phonological and metrical variables by Prosodic (Heuser, Falk, and Anttila 2017).

<u>Original text (O)</u>		<u>Randomized text (R)</u>		<u>Scrambled text (S)</u>	
<u>Words</u>	<u>Parts of speech</u>	<u>Words</u>	<u>Parts of speech</u>	<u>Words</u>	<u>Parts of speech</u>
For	IN	of	IN	by	IN
my	PRP\$	my	PRP\$	fittest	JJS
religion	NN	sand	NN	terms	NNS
,	,	,	,	,	,
though	RB	still	RB	it	PRP
there	EX	there	EX	Christ	NNP
be	VBP	doth	VBP	which	WDT
several	JJ	reserved	JJ	their	PRP\$
circumstances	NNS	urns	NNS	good	JJ
that	WDT	which	WDT	by	IN
might	MD	would	MD	less	CC
persuade	VB	make	VB	past	NN
the	DT	the	DT	those	DT
world	NN	day	NN	easy	JJ
I	PRP	him	PRP	not	RB
have	VBP	have	VBP	in	IN
none	NN	subject	NN	parents	NNS
at	IN	from	IN	of	IN
all	DT	the	DT	sorrow	NN

Table 1: The first twenty words in Browne’s *Religio Medici* are shown in their original in the leftmost column, along with the part of speech for those words in the adjacent column. In the middle two columns, the words from the original have been randomized, but part of speech is preserved. In the rightmost two columns, the words from the original have been scrambled, ignoring part of speech.

### 3. Results

One of our goals is to find out whether standard constraints posited by phonologists and metricists are useful for distinguishing rhythmic varieties of prose, potentially along the metrical, antimetrical, anti-antimetrical, and ametrical lines hypothesized in our introduction. This test requires a framework for automatic scansion which we have in Prosodic. As is often the case in research, trying to give an explicit answer to a specific question leads to new discoveries unrelated to the original question, moving the research along. We will see that many generalizations proposed by prosodists like Saintsbury based on intuitions stand up to empirical scrutiny in terms of corpora. We will also see that formal tools (automatic scansion, statistical methods) are useful because they allow us to reach generalizations not accessible to earlier researchers and that shed light on the structure of texts as well as on the structure of language.

We present here three small experiments. First, we test for Weak Antimetricality by looking within each author for stretches of language that conform to the constraints defining blank verse, assessing whether such stretches of iambic pentameter appear more often than we would expect by chance. Second, we test for Strong Antimetricality by conducting a broader investigation into the metricality of our authors, comparing which authors most consistently observe the metrical constraints in (9). Finally, we apply our test of Strong Antimetricality to a large digital corpus, “distant-reading” the rhythmic history of verse and prose.

#### 3.1. Weak Antimetricality

As the Self-Taught Man makes clear in Sartre’s *Nausea*, to avoid stretches of language that conform to a canonical metrical template such as hexameter (or, in the case of English, pentameter), was historically an explicit principle of prose style. A well-known aspect of Cicero’s rhetoric was the ending of sentences with a “clausula,” a rhythmic pattern defined precisely by its non-alternating rhythm, or its antimetricality. Because prose writers may have *consciously* avoided such instances of metricality, we call this experiment a test of “weak antimetricality,” as opposed to a stronger form in which prose writers may have avoided metrical patterning more broadly and unconsciously.

For this experiment, we gather all ten-syllable stretches of words occurring just before a punctuation break, across all six authors. Each of these *potential* pentameters is then scanned metrically in order to determine whether it is, in fact, a *valid* line of iambic pentameter. We

consider as valid any line that has five strong positions (pentameter); is rising in rhythm (iambic); and does not violate \*W/PEAK (9a), W-RESOLUTION (9d), and F-RESOLUTION (9e), all of which constraints are normally thought mandatory for the iambic pentameter line. We then perform this same process for a series of *randomized* texts (see Table 1 above for an example). We copy examples of valid and invalid pentameters below, from both the original texts (Table 2) and randomized texts (Table 3).

	<b>Original text</b>	
	<i>Valid Pentameter</i>	<i>Invalid Pentameter</i>
<b>Shakespeare</b>	And burn the long-lived phoenix in her blood	bounteous gift thou shouldst in bounty cherish
<b>Browne</b>	A <u>power</u> to <u>move</u> the <u>body</u> <u>it</u> <u>informs</u>	bad cause is but constancy in a good
<b>Dickens</b>	or wooden causeway we had strolled upon	after a few seconds of such repose
<b>Pater</b>	for after passing thus from hand to hand	in its quaint fragments of early French prose
<b>Ruskin</b>	And law of aspect hewn of half her aisles	Roman epithet above its houses
<b>Dibble</b>	uneven edges of the joint appear	wiped considerably more easily

Table 2: From the original texts for each author, an example valid pentameter line (in the left column), and an example invalid pentameter line (in the right column). All examples come from 10-syllable stretches of words immediately preceding a punctuation mark.

	<b>Randomized text</b>	
	<i>Valid Pentameter</i>	<i>Invalid Pentameter</i>
<b>Shakespeare</b>	uncertain outward deeds in thy twain moon	those love yet reason at your example
<b>Browne</b>	to the true wisdom of community	their visible little memorandum
<b>Dickens</b>	seemed I upon his pleasure with it had	the subjects affronted experience
<b>Pater</b>	the shadowy creation of the flesh	own mirthfulness forgets upon dazzled
<b>Ruskin</b>	hand is far more laboriously cut	central epoch of the life of Venice
<b>Dibble</b>	with joint and lead to use and turn this trench	the foundation dug in having the joint



Table 3: From *the randomized texts* for each author, an example valid pentameter line (in the left column), and an example invalid pentameter line (in the right column). All examples come from 10-syllable stretches of words immediately preceding a punctuation mark.

How often do we encounter these pentameter lines? For Shakespeare, 75% of 10-syllable phrases preceding punctuation were deemed valid pentameter; for Dibble, that number drops to 32%. Sorting the authors by these frequencies of valid pentameter lines, we obtain the following ranking of the authors as an index of Weak Antimetricality:

(17) Shakespeare > Dickens > Dibble > Browne > Pater > Ruskin

However, we should expect some of these 10-syllable phrases to be valid pentameter by chance. This is why we compare these percentages in the actual texts to those drawn from a series (25) of randomized texts, in order to compare how often we observe valid pentameter lines to occur in the authors' actual texts, with how often we expect them to occur by chance. In Figure 1 below, we visualize both the observed and expected frequencies. Figure 1 shows that, unsurprisingly, Shakespeare's sonnets include many more lines of iambic pentameter than we would expect by chance. So far, so good: although unsurprising, this result lends credence to the remaining findings. Several of these findings support the antimetricality hypothesis, although a couple do not.

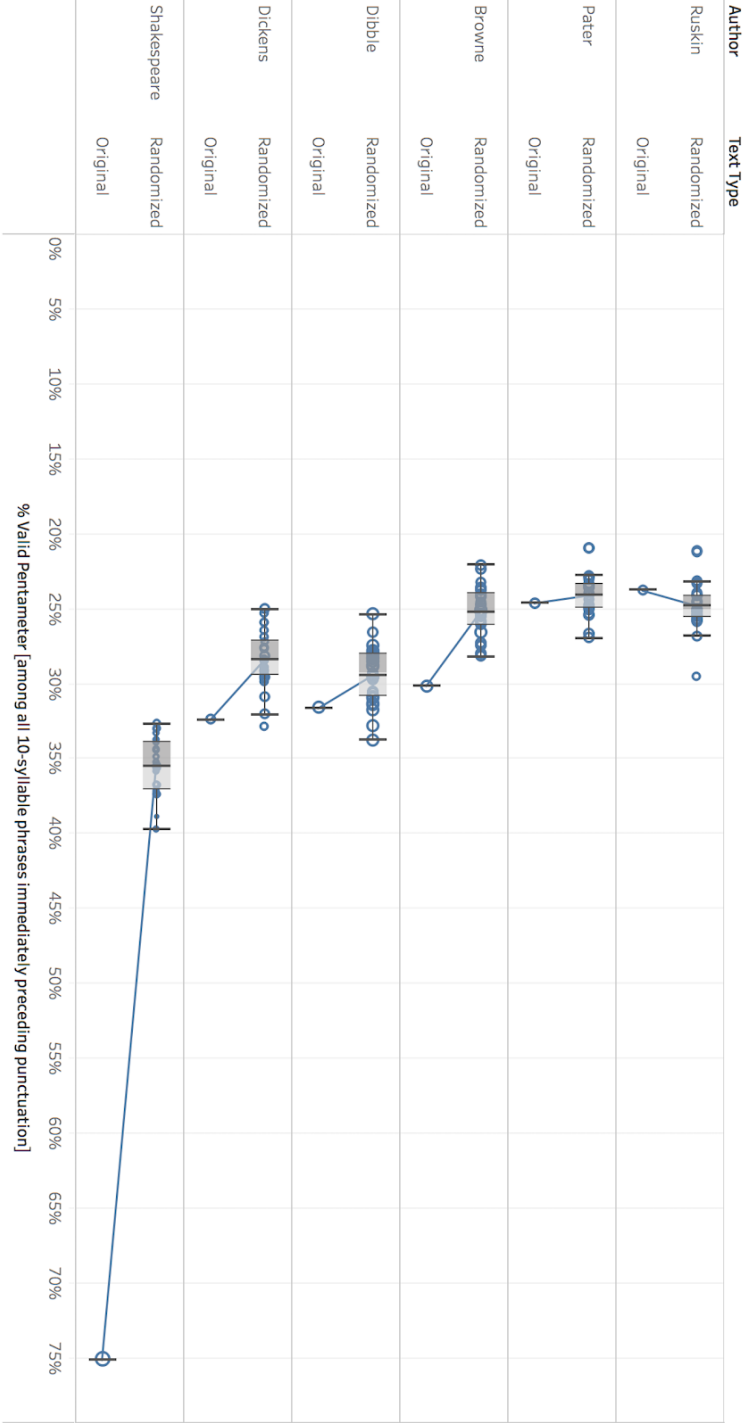
In support of our hypothesis, Dickens, thought to be an author flirting with meter, is the next most likely author after Shakespeare to include valid iambic pentameter lines in his prose. In addition, the frequency with which iambic pentameter lines appear in Dickens' *David Copperfield* is significantly greater than the frequency with which they would appear by chance. Also in support of our hypothesis is that our two anti-metrical authors, Browne and Pater, include relatively fewer iambic pentameter lines than does our ametrical author, Dibble, author of a plumbing manual, which we take to have no metrical agenda. Ruskin, in spite of his heightened passages like (5), on the basis of the bulk of his prose clocks in with overall the fewest iambic pentameter lines among all six authors, even fewer than would be expected by chance, the only author to do so.

One result, however, is inconsistent with our hypothesis. What does it mean that Browne, who for us and Saintsbury is an example of antimetricality, includes significantly more iambic pentameter lines than would be expected by chance?

Ultimately, then, the results of this first experiment are suggestive, but inconclusive. This is not unexpected. The relative frequency with which iambic pentameter sequences end phrases is, we think, an interesting indicator of what we have called Weak Antimetricality, a specific and well-known manifestation of meter within prose that many authors sought explicitly to avoid. The extent to which prose avoids the metrical patterning of verse is a phenomenon broader than the incidence of iambic pentameter, a particular meter. The suggestive but inconclusive outcome

of this experiment, therefore, lends further interest to the following section's more nuanced and multidimensional model of the relative metricality of prose.

Accidental Pentameter Index



**Figure 1.** Along the x-axis is displayed the percentage of *potential* iambic pentameter lines (i.e. 10-syllable phrases occurring just before punctuation marks) that, metrically, are in fact *valid* iambic pentameter lines. This percentage is recorded for the original text, on the second row for each author; and also for 25 randomized texts, on the first row. The percentage of valid pentameter in the randomized texts serves as a baseline expectation, the percentage of valid pentameter we might expect to see simply by virtue of the texts' diction, without any conscious intent regarding the sequencing of these words. The line between the median for the randomized texts' percentages, and the actual text's percentage, records the distance between the baseline expectation and the observed result.

### 3.2 Strong Antimetricality

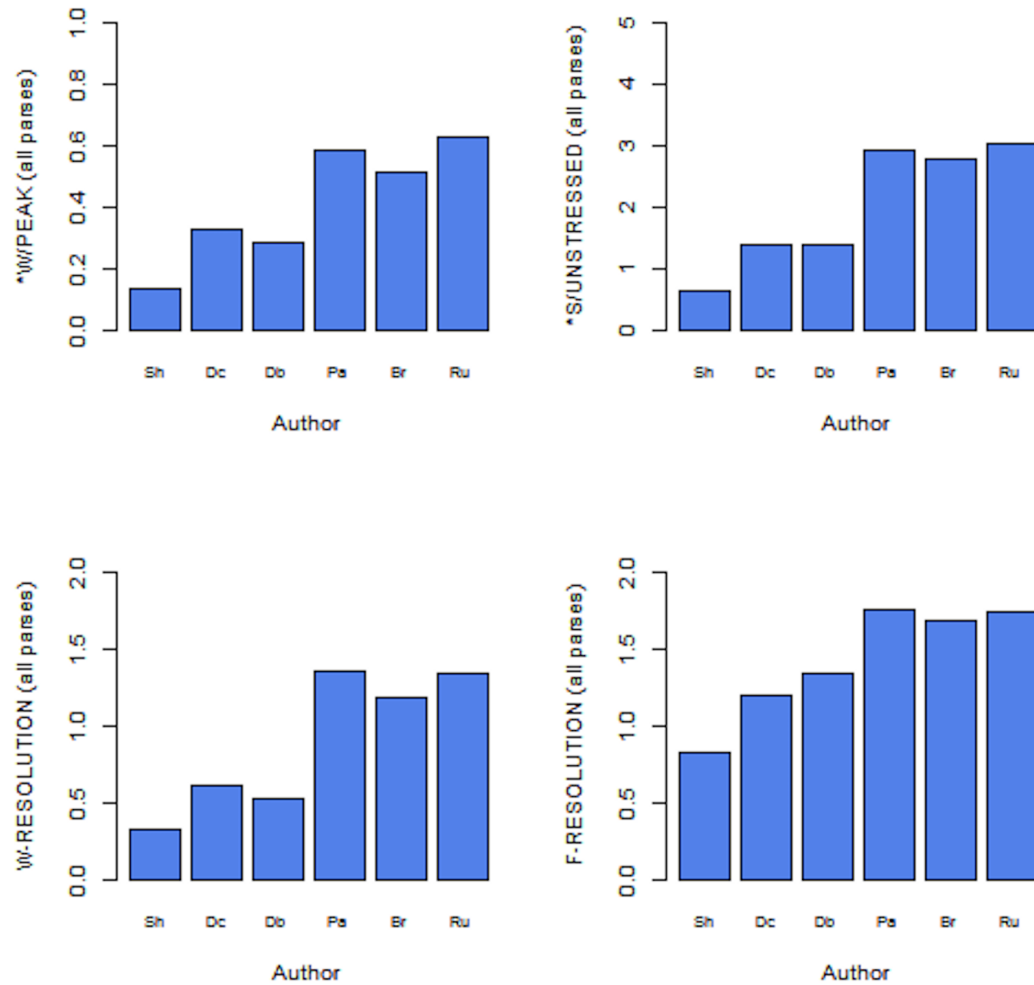
As mentioned in our introduction, Strong Antimetricality hypothesizes that prose avoids all forms of metrical patterning, frustrating the expectation of alternation in a search for rhythmic complexity and variety. Because this makes antimetricality independent of any particular meter, it is necessary to parse each line according to a wide range of possible metrical templates. As explained in section 2.2, the automatic scansion tool, Prosodic, scans lines by applying every logically possible scansion, evaluating each by how well it fits the phonology of the line. Specifically, each scansion is evaluated by how often it violates a metrical constraint, a rule regulating the correspondence of metrical template to phonological pattern.

By applying the five metrical constraints laid out in section 2.1, Prosodic parsed 9,000 pseudo-lines (i.e. 5-word fragments) from each of our six authors. By measuring how often these pseudo-lines violate the metrical constraints, then, we effectively measure the extent to which snippets of language resist metrical scansion. Recall that such violation is not unusual, even in verse. To conform perfectly to a set of metrical constraints is the stuff of nursery rhymes, and even these occasionally vary their rhythm against their meter. The *frequency* with which such violations occur, however, indexes the relative degree of friction or resistance between a text and a given metrical template. If the well-mannered meters of nursery rhymes lie at one end of this spectrum, only infrequently violating metrical constraints, then complex form of prose to lie at the other end, violating metrical constraints so frequently that it seems no longer to make sense to apply them. Nevertheless, to apply metrical constraints to prose as well as verse is to place both forms of language on a rhythmic continuum with each other. We call this continuum of metrical constraint violation METRICAL FRICTION, and consider it one promising measurement of Strong Antimetricality, or the pervasive avoidance of metrical patterning.

Ranking authors according to the frequency with which they violate specific metrical constraints arranges them more or less in the same way, regardless of the constraint, picking out Shakespeare as the most metrical author (see Figure 2). There seem to be three groups of authors. Shakespeare is metrical by all four constraints. The second place is held by Dickens and Dibble, in some order, and the third place by Pater, Browne, and Ruskin. The fact that the authors are placed in more or less the same order by each metrical constraint suggests that sum of their violations should be a sensible measure of antimetricality.<sup>14</sup>

---

<sup>14</sup> These results were corroborated through a statistical model. In a linear regression model with the total number of violations as the dependent variable and author and the number of syllables as predictors, the authors are divided into two groups. Shakespeare, Dickens, Browne, and Ruskin all have significantly fewer violations than Dibble (the baseline) and Pater. This is broadly consistent with what we see in the visualizations.



**Figure 2:** Average number of individual constraint violations per line for each of the six authors. These four continua of metrical violation typically run from Shakespeare (Sh), with the least metrical friction to Dickens (Dc) and Dibble (Db), with more metrical friction to Pater (Pa), Browne (Br), and Ruskin (Ru), with the highest degree of metrical friction.

Another measure of antimetricity follows from the first: the number of plausible scansions for a line, or what we call METRICAL UNCERTAINTY. As demonstrated in section 2.2, a line that violates a number of metrical constraints can often be scanned in a number of distinct ways, all of which represent plausible interpretations of the line.

The architecture of Prosodic is adopted from Optimality Theory (Prince and Smolensky 1993/2004). Prosodic starts by considering all the logically possible candidate scansions and picks the best one among them as optimal. The following table illustrates this for the line *thou that art now the world's fresh ornament* given three candidate scansions: (a), (b), and (c). We consider two constraints: HOPKINS and W-RESOLUTION. The numbers stand for constraint violations.

<i>thou that art now the world's fresh ornament</i>	HOPKINS	W-RESOL
(a) thou THAT art NOW the WORLD'S fresh OR.na. <u>MENT</u>	1	
(b) thou THAT art NOW the WORLD's fresh <u>OR.NA</u> . <u>MENT</u>	1	2
(c) thou THAT art NOW the WORLD'S fresh OR. <u>na.ment</u>		1

HOPKINS assigns a violation for each unstressed syllable in a strong position. It is violated by the syllable MENT in candidate scansions (a) and (b). W-RESOLUTION assigns a violation for each disyllabic position within a word where the first syllable is not light and stressed. It is violated by the disyllabic positions OR.NA and NA.MENT in (b), hence two violations, as well as by *na.ment* in (c). Prosodic now applies a principle by which a candidate scansion is rejected if it has a superset of violations compared to some other candidate scansion. Scansion (b) is rejected on two independent grounds: it incurs all the violations that (a) incurs plus two additional violations of W-RESOLUTION, and similarly incurs all the violations that (c) incurs plus one additional violation of HOPKINS and another violation of W-RESOLUTION. In the technical jargon of Optimality Theory, doomed candidates like (b) are called HARMONICALLY BOUNDED.<sup>15</sup>

Harmonic bounding leaves us with two competing scansions, (a) and (c), of which Prosodic selects the one with the fewest violations. Since both (a) and (c) have exactly one violation Prosodic accepts both. The number of scansions that survive past harmonic bounding is interesting as it provides valuable information about the metricality of the line. A large number of survivors indicates that there is no single scansion that clearly beats the others, suggesting that the line is metrically ambiguous. This corresponds to a familiar experience in reading verse: finding the meter usually takes a few lines of trial and error before one finds the pulse of the text, and sometimes there is no pulse to be found, suggesting that there are too many competing

<sup>15</sup> The harmonically bounded scansion (b) was constructed by hand as Prosodic currently only outputs the harmonically unbounded scansions.

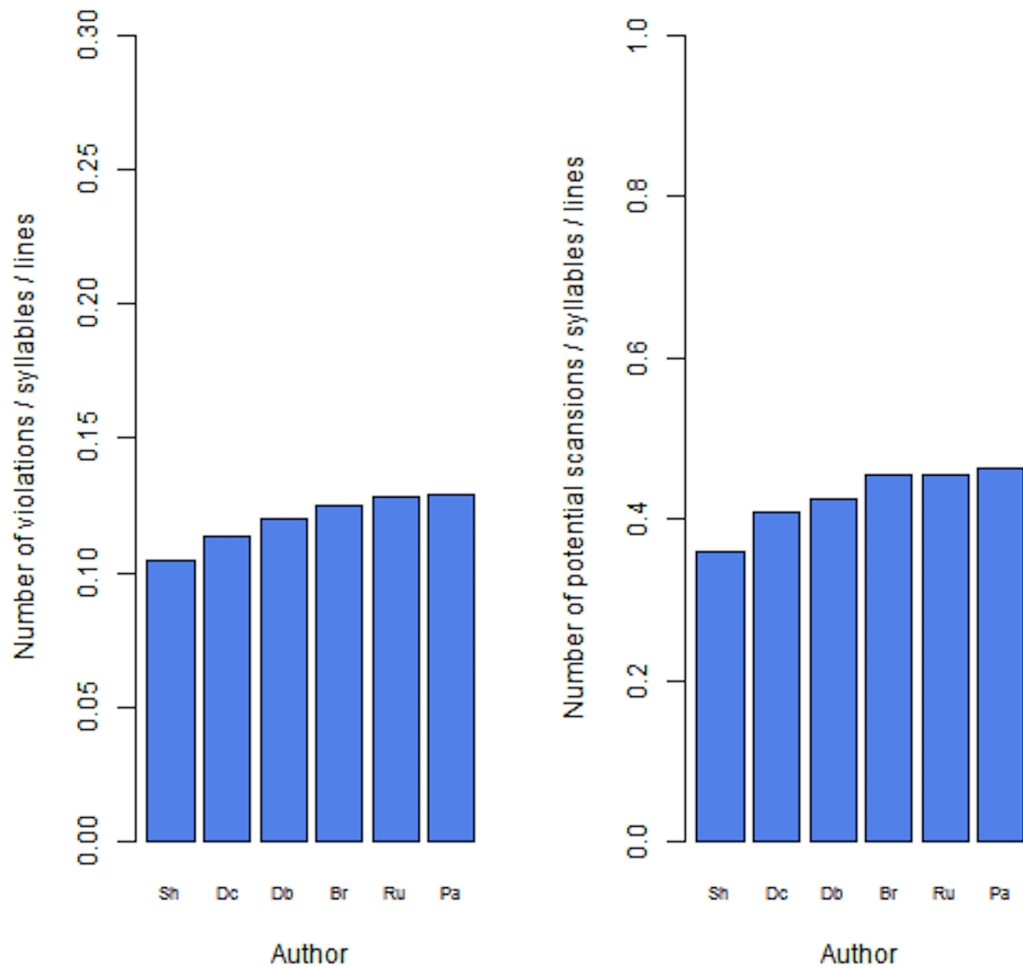
scansions to consider. The number of surviving scansions thus provides an intuitive measure of metricality: few survivors means that the meter is easy to find; a single survivor means that it is very easy to find. In contrast, many survivors means there are many intuitively plausible ways to scan the sentence. This is what we call METRICAL UNCERTAINTY: the more surviving scansions, the less certain we can be which one of them is correct.

In order to compare this measure to the number of violations measure, we show both in Figure 3. In both plots, the raw frequency is divided by the number of syllables in the line and the number of lines in the author's text to make the scores comparable across authors. Both measures rank the authors in the same order of (anti-)metricality.

(18) Shakespeare > Dickens > Dibble > Browne > Ruskin > Pater

This ranking mirrors that given in (17) regarding accidental blank verse as an indicator of Weak Antimetricality, with the only exception that Ruskin's and Pater's positions are inverted. Similarly, then, this result suggests that the hypothesis of Strong Antimetricality may be broadly correct, with a few exceptions. Compared to the ametrical Dibble, the antimetrical Browne and Pater are less metrical, and the anti-antimetrical Dickens is more metrical. This is consistent with the visualizations we have seen so far and is also broadly supported by our preliminary regression models. Note that Ruskin falls on the antimetrical side of the spectrum. By both weak and strong measures of antimetricality, he is less metrical than Dibble.

In sum, both measures of metricality rank our six authors according to the order predicted by the Antimetricality Hypothesis, with the reservation that it does not catch the metrical flights embedded in Ruskin's apparently otherwise antimetrical prose. In the next and final results section, we scale up the experiments developed here and in the previous section, applying their measures of antimetricality to thousands of works of prose and verse across literary history.



**Figure 3:** On the left, the number of constraint violations, or METRICAL FRICTION, in each author; on the right, the number of plausible scansions, or METRICAL UNCERTAINTY.

### 3.3 Big Metrics

So far we've looked at six authors meant to represent four different metrical positions: the metrical (Shakespeare), opposed by the antimetrical (Browne and Pater), in turn opposed by the anti-antimetrical (Dickens and Ruskin)—and all of whom ignored by the ametrical (Dibble). Admittedly, it's asking quite a lot of our authors to represent an entire form of (anti-)metricality, rather than, say, the metrical tendencies of the specific text we chose, or perhaps the authors'



personal metrical style, or the metrical tendencies of their historical period. In this section, then, we scale up the experiment to 1,163 authors and 2,903 works of verse, prose non-fiction, and prose fiction published across 1700-1900.<sup>16</sup> We apply to each of these texts the same measures of antimetricality developed in the previous sections, taking from each text 500 5-word fragments and parsing them according to the procedures outlined above. With thousands of texts from the eighteenth and nineteenth centuries, our corpus covers the period during which the novel displaces poetry as the dominant literary form. At this larger scale, and over this crucial period, changes in these measurements should indicate whether, as we hypothesize, prose becomes less antimetrical as verse becomes less dominant.

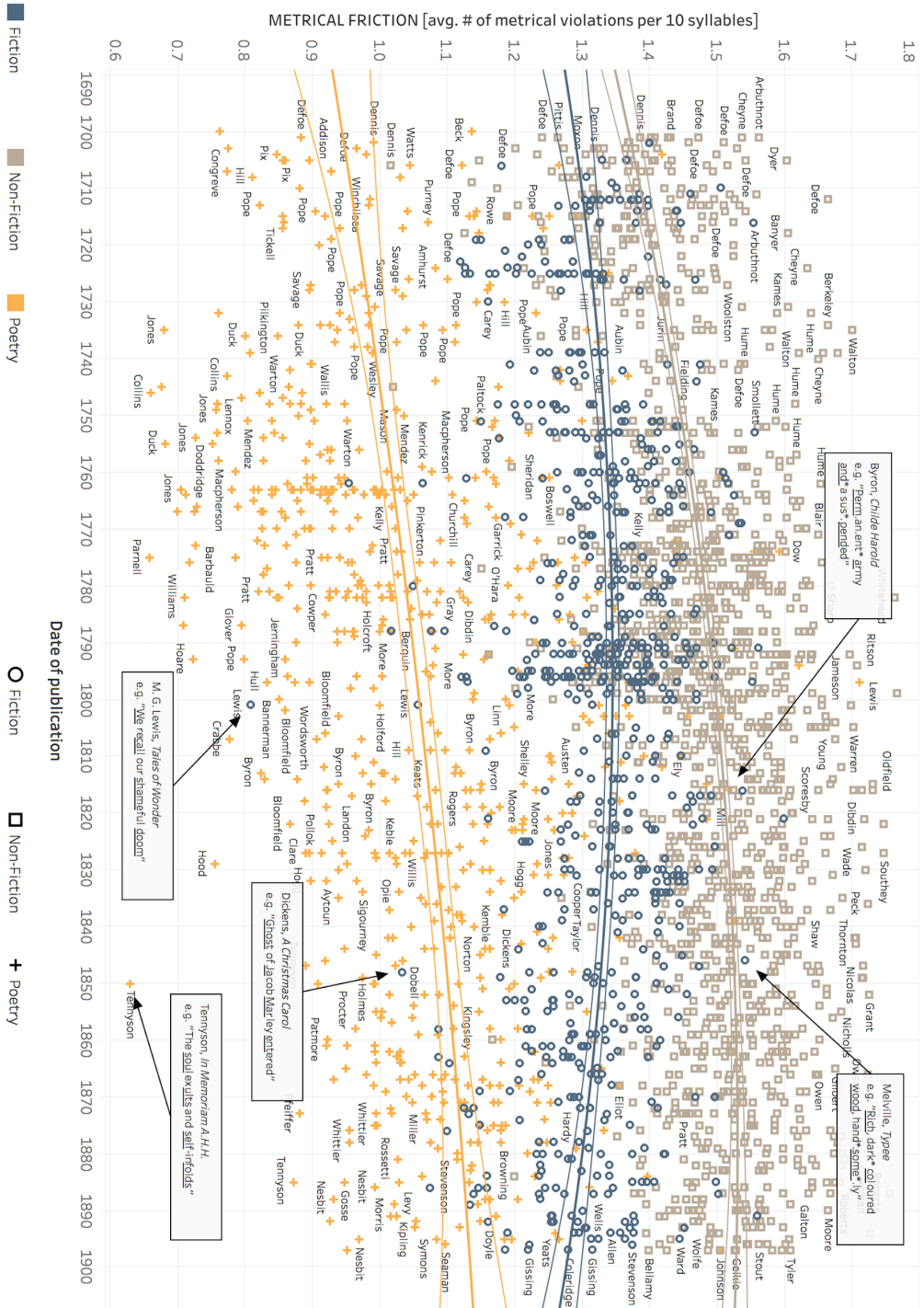
Figure 4 show that this may be the case. Plotting Metrical Friction across these 3,000 texts produces, we believe, a new literary history, a history of antimetricality in verse and prose across the eighteenth and nineteenth centuries.<sup>17</sup> This is a complex history, involving simultaneous and intersecting trends across poetry, prose non-fiction, and fiction. From the figure, for example, we can clearly see that poetry (in orange crosses) is *increasing in metrical friction* across this period. Insofar as metrical friction is one measurement of antimetricality, then poetry itself is becoming more antimetrical from 1700 to 1900. Such a transformation toward the antimetrical within verse is not, of course, unexpected. The heroic couplets and neoclassical metrics of Alexander Pope give way, as we know, to Browning's ranging monologues, Hopkins' sprung rhythm, Robert Frost's "loose iambics," and ultimately the total rejection of meter in the twentieth century.<sup>18</sup> However, these data show that the poetry's disavowal of meter was a smooth and continuous historical process, not a sudden revolution.

---

<sup>16</sup> Unfortunately, we do not yet have data for the seventeenth or the twentieth century. Both of these periods play an important role in the Antimetricality Hypothesis, with perhaps an especial emphasis on the twentieth century, which sees the rise of free verse within poetry and the (hypothesized) rise of ametricality within prose. We are currently working to add texts from both periods in order to round out this narrative.

<sup>17</sup> Recall that Metrical Friction, discussed above in section 3.2, indexes a text's resistance to metrical scansion by measuring the frequency with which its actual phonological patterning resists even its best-fitting metrical template.

<sup>18</sup> For a recent account of Frost's metrical "looseness," see Ben Glaser, "Modernist Scansion: Robert Frost's Loose Iambics," *ELH* 83.2 (2016), 603-31. Although Glaser makes his case independent of Prosodic and its assumptions about meter, Glaser applies Prosodic to Frost in a footnote to find that its output substantiates his argument.



**Figure 4.** Metrical Friction calculated across 2,903 works of verse, prose non-fiction, and prose fiction published 1700-1900. Metrical Friction indexes a text's resistance to scansion through measuring the frequency with which it violates the metrical templates imposed onto its actual phonological patterns (e.g. stress). Trend lines show a binomial regression for each genre, all of which are statistically significant ( $p$ -value < 0.0001).

More importantly, perhaps, these data also show that poetry's growing antimetricality intersects with related metrical trends in fictional and non-fictional prose. At the beginning of the eighteenth century, these two prose genres are metrically indistinct, and both far less metrical than verse. Although Defoe's poem "Ye True-Born Englishman Proceed," for example, is much more metrical than his prose, his novel *Moll Flanders* (1723) is no more or less metrical than his "The Chimera: or the French Way of Paying National Debts Laid open" (1720).<sup>19</sup> However, over the eighteenth and nineteenth centuries, fictional and non-fictional prose undergo a metrical separation. By the end of the nineteenth century, fictional prose is metrically closer to poetry than it is to non-fiction. At this point, novels like H. Rider Haggard's *She* (1887), with perfectly metrical pseudo-lines like "as fire faintly shines within" and "utter darkness in the bowels," are in fact *more* metrical than a handful of volumes of poetry—like Whitman's *Poems* (1868), full of such uncertain pseudo-lines as "traversing paved avenues, amid temples" and "nothing endures but personal qualities." If verse becomes *less* metrical over the eighteenth and nineteenth centuries, these data reveal that, over the same period, fiction becomes *more* metrical—a fact that, as far as we know, is new to literary history.

However, that prose fiction becomes more metrical across this period is predicted by the Antimetricality Hypothesis. According to that hypothesis, prose drops its oppositional stance toward meter—its antimetricality—to the extent that metrical verse is no longer the dominant literary form. And yet, although we see this drop in prose fiction, we see quite the reverse in prose non-fiction, which is only further increasing its antimetricality. This complication suggests at least two possible interpretations of the data within the Antimetricality Hypothesis. As mentioned above, in the early eighteenth century metricality divides verse from prose, but does not distinguish between fictional and non-fictional prose. Subsequently, however, through a simultaneous fall of metricality in verse and a rise of metricality in prose, the rhythmic boundary has shifted to lie between the "literary" (fiction and poetry) and the "non-literary" (non-fiction)—where the word "literary" here has its nineteenth-century sense as imaginative writing in prose or verse. Accordingly, one interpretation of the Antimetricality Hypothesis is that not prose style generally, but specifically *literary* prose becomes less antimetrical. On this view, metricality may have increasingly become a signifier for the literary, rather than the poetic, as verse falls from dominance and its metricality converges with that of fiction.

Another, related but independent interpretation of the Antimetricality Hypothesis is that specifically *fictional*, rather than "literary," prose drops its stance of antimetricality over this period. It should be noted that verse does not so much fall from dominance as it is displaced by the specific cultural form of the novel. The novel is also widely thought to incorporate within itself a range of outlying discursive modes and genres, like the biography, epistolary, diary, and newspaper. On this interpretation, then, once the novel had displaced verse from dominance, it

---

<sup>19</sup> Defoe's poem "Ye True-Born Englishman Proceed" has a Metrical Friction score of 0.883; his novel *Moll Flanders*, a score of 1.398; and his pamphlet "The Chimera," a score of 1.403.

became free to incorporate what was then a crucial formal signifier of verse—meter—into its novelistic form. This view has the advantage of simultaneously explaining why verse may have found itself “flirting with prose,” gradually dropping its anti-prosaic stance. That Elizabeth Barrett Browning presented her epic poem, *Aurora Leigh* (1856), as a “novel in verse” speaks to the literary market-pressures conditioning the publication of verse within a novel-saturated society. These pressures may help explain poetry’s formal drift toward the antimetricality of prose, as poetry responds to its displacement from dominance by attempting to incorporate the formal attractions of its disposessor, the novel—perhaps especially the novel’s facility with narrative realism and what are sometimes called the “prosaic” aspects of life, but perhaps also secondarily with the novel’s comparative antimetricality, its *rhythmic* prosaicness. Although none of these changes is sudden or revolutionary, Wordsworth may have been articulating these slow-moving cultural forces when he argued in the 1802 preface to the *Lyrical Ballads* that “there neither is, nor can be any essential difference between the language [i.e. the diction] of prose and metrical composition,” and accused eighteenth-century critics and poets with “attempt[ing] to widen the space of separation betwixt Prose and Metrical composition.” For Wordsworth, however, it is exclusively *meter* that legitimately separates verse from prose<sup>20</sup>; it is therefore ironic that, in the century following his manifesto, we find not only the linguistic<sup>21</sup> but also the *metrical* “space of separation” between verse and fictional prose collapsing.

#### 4. Conclusion

This project, involving a team of linguists and literary scholars, is a study of literary history using methods from corpus linguistics and theories from metrical phonology, a branch of linguistics within which Shakespeare, Milton, and Hopkins present central theoretical challenges. The intellectual range in the interests and findings produced by this project is as wide-ranging as its components. Along the way, we studied phenomena ranging from the role of syntax in text randomization, and the role of sonority and syllable weight in predicting metricality; to the history of specific meters and clausulae, and even the history of metricality itself across the modern era. These pages are a first articulation of some of our findings, but they remain provisional. We continue to explore the questions central to all of the various components of our work: what is rhythm in verse, what is rhythm in prose, and how can we better relate them?

---

<sup>20</sup> On the virtues of metricality in verse, Wordsworth writes in the same preface: “And it would be a most easy task to prove to him, that not only the language of a large portion of every good poem, even of the most elevated character, must necessarily, *except with reference to the metre*, in no respect differ from that of good prose, but likewise that some of the most interesting parts of the best poems will be found to be strictly the language of prose, when prose is well written”; “and, *if metre be superadded thereto* [to the ‘language really spoken by men’], I believe that a dissimilitude will be produced altogether sufficient for the gratification of a rational mind.”

<sup>21</sup> In a foundational digital humanities experiment, Ted Underwood and Jordan Sellars have demonstrated that, between 1750 and 1900, the language of poetry and fiction gradually move in the Wordsworthian “direction of old words that would appear plain, common, and universal,” thereby separating from the language of non-fiction prose. See Underwood and Sellars, “The Emergence of Literary Diction,” *Journal of Digital Humanities* 1.2 (2012).

## References

- ANTTILA, ARTO & RYAN HEUSER. 2016. Phonological and metrical variation across genres, *Proceedings of 2015 Annual Meeting on Phonology*, Linguistic Society of America, Washington DC.  
<https://journals.linguisticsociety.org/proceedings/index.php/amphonology/article/view/3679>
- BLUMENFELD, LEV. 2015. Meter as faithfulness, *Natural Language and Linguistic Theory*, 33(1), 79-125.
- GODZICH, WLAD & JEFFREY KITTAY. 1987. *The Emergence of Prose: An Essay in Prosaics*. University of Minnesota Press.
- GOLSTON, CHRIS. 1998. Constraint-based metrics, *Natural Language and Linguistic Theory* 16(4), 719–770.
- GOLSTON, CHRIS & TOMAS RIAD. 2000. The phonology of classical Greek meter. *Linguistics* 38(1), 99–167.
- GOLSTON, CHRIS & TOMAS RIAD. 2005. The phonology of Greek lyric meter. *Journal of Linguistics* 41(1), 77–115.
- 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 & PAUL KIPARSKY. 1996. A parametric theory of poetic meter. *Language* 72: 287– 335.
- HANSON, KRISTIN & PAUL KIPARSKY. 1997. The Nature of Verse and its Consequences for the Mixed Form. In J. Harris and T. Ziolkowski (eds.) *Prosimetrum*. Cambridge: Brewer, 1997.
- HAYES, BRUCE, COLIN WILSON, & ANNE SHISHKO. 2012. Maxent grammars for the metrics of Shakespeare and Milton, *Language* 88(4), 691-731.
- 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>.

JAKOBSON, ROMAN. 1935a. The Dominant. Reprinted in *Language and literature*, Eds. Krystyna Pomorska and Stephen Rudy. Cambridge: The Belknap Press of Harvard University Press.

JAKOBSON, ROMAN. 1935b. Marginal Notes on the Prose of the Poet Pasternak. Translated and reprinted in *Language and literature*, Eds. Krystyna Pomorska and Stephen Rudy. Cambridge: The Belknap Press of Harvard University Press.

JAKOBSON, ROMAN. 1941. Kindersprache, Aphasie, und allgemeine Sprachgesetze. The Hague: Mouton. Selected Writings, Vol. 1.

KIPARSKY, PAUL. 1977. The Rhythmic Structure of English Verse. *Linguistic Inquiry* 8: 189-247.

KIPARSKY, PAUL. 2006. A modular metrics for folk verse. In *Formal approaches to poetry*, eds. Bezalel Elan Dresher and Nila Friedberg, 7–49. Mouton: Hague.

KIPARSKY, PAUL. 2018. Indo-European Origins of the Greek Hexameter. In Gunkel, Dieter & Olav Hackstein (eds.) *Sprache und Metrik*. Leiden: Brill.  
<https://web.stanford.edu/~kiparsky/Papers/hexameter.pdf>

PRINCE, ALAN. 1989. Metrical forms, in Paul Kiparsky and Gilbert Youmans (eds), *Phonetics and Phonology I: Rhythm and Meter*, Academic Press, San Diego, pp. 45-80.

PRINCE, ALAN AND PAUL SMOLENSKY. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*, Blackwell Publishing, Malden, Massachusetts.

PURNELLE, GÉRALD. "Excusez-moi, monsieur": Les alexandrins dans *La Nausée*. *Poétique* 2005/4 (n° 144). <https://www.cairn.info/revue-poetique-2005-4-page-457.htm#re5no5>

SAINTSBURY, GEORGE. 1912. *A History of English Prose Rhythm*. First Indiana University Press edition 1965, first published in 1912.

SPIEGEL, GABRIELLE M. 1993, *Romancing the Past: The Rise of Vernacular Prose: Historiography in Thirteenth-Century France*. University of California Press, Berkeley.  
<http://ark.cdlib.org/ark:/13030/ft209nb0nm/>

STEELE, TIMOTHY. 1999. *All the Fun's in How You Say a Thing*, Ohio University Press,

Athens.

WEIDE, R. L. 1998. The CMU Pronouncing Dictionary, release 0.6 [syllabification, stress, and weight annotation added by Michael Speriosu],  
<http://web.stanford.edu/~anttila/research/sylcmu.txt>.