Stress-weight and stress-tone interaction in South Slavic folk meter

Abstract

Bosnian/Croatian/Montenegrin/Serbian exhibits three types of prominence in syllables, namely, stress, weight, and tone. All three types, we demonstrate, are regulated by its poetic meter, of which we focus on the epic decasyllable, the most widely used folk verse form. Specifically, we show that weight and tone are regulated in stressed syllables, with three main implications for metrical theory and typology: First, a single meter can impose constraints not just on multiple types of prominence, but on multiple interactions of those types (here, stress-tone and stress-weight). Second, one of these interactions, stress-tone, has not previously been documented for any meter. Third, this is the first demonstration of an Indo-European meter being sensitive to tone. Using Maxent and other models, we show that the poets productively manipulate these combinations of prominence beyond their baseline rates. Moreover, multiple violations of the constraints interact cumulatively, favoring Harmonic Grammar (e.g. Maxent) as a theory of phonology.

Keywords: metrics, epic decasyllable, stress, weight, tone, Harmonic Grammar, Maxent

Word count: 11,913 (excluding references)

1 Introduction

Across languages, stress interacts with tone and syllable weight. Stress can be attracted by High tone, and Highs can be avoided in unstressed positions (Goldsmith, 1987; Bickmore, 1995; Yip, 2002; de Lacy, 2002; Gordon, 2023). Likewise, stress can be attracted to heavy syllables, or avoided on lights (Kager, 1989; Prince, 1990; Prince and Smolensky, 2004; Hayes, 1995; Gordon, 2006; Ryan, 2016). This article explores how stress, tone, and weight interact in poetic meter, where they map onto yet another dimension of prominence—metrical strength. Our study is couched in generative metrics (Halle and Keyser, 1966, 1971; Kiparsky, 1975, 1977 *et seq.*), and more specifically, the modular templatic approach (*ibid.*), which views meter as an abstract template of alternating strong and weak positions organized into larger constituents such as feet, cola, hemistichs, and lines. The grammar of the meter then governs the correspondence between this metrical structure and linguistic material.

Many meters impose restrictions on only a single prosodic property. In ACCENTUAL (SYLLABOTONIC) meters, for instance, stressed syllables are favored in strong positions and/or avoided in weak ones (e.g. English; Hayes et al. 2012). QUANTITATIVE meters (e.g. Ancient Greek), by contrast, regulate syllable weight: Heavy syllables are favored in strong positions and/or avoided in weak ones. Lastly, some meters (e.g. Vietnamese) are TONE-BASED, mapping tone onto the metrical template (see Hanson and Kiparsky, 1996; Fabb and Halle, 2008; Blumenfeld, 2016; Kiparsky, 2020 for overviews of metrical typology).

Additionally, some meters are HYBRID, regulating multiple prosodic features simultaneously. The Latin dactylic hexameter, for instance, is a hybrid meter featuring the INDEPENDENT MAPPING of two properties, namely, weight and stress (Ryan, 2017). All strong positions must contain a heavy, to satisfy (1a). Moreover, strong positions in the cadence (final two feet of the line) are virtually always stressed, in response to (1b). Crucially, weight and stress mapping are orthogonal: Neither constraint in (1) refers to both weight and stress (*ibid.*). (As Ryan (2017) argues, the regulation of stress in the cadence cannot be explained away by constraints on weight.)

(1) a. STRONG⇒HEAVY
A strong position must contain a heavy syllable.

b. STRONG⇒STRESS$_{\text{cadence}}$
A strong position in the cadence must contain a stressed syllable.

A hybrid meter might also exhibit INTERACTIVE MAPPING, whereby multiple prosodic dimensions are regulated jointly, by the same constraint. For example, the Finnish Kalevala meter regulates weight, but only does so strictly for syllables with primary stress, as Ryan (2017) analyzes with the constraint in (2). Among unstressed and secondary-stressed syllables, weight is less strictly regulated.

(2) STRESS⇒(STRONG⇔HEAVY)
If a syllable is primary-stressed and assigned to a strong position, it must be heavy; if primary-stressed and heavy, it must be assigned to a strong position.

In this article, we show that the epic decasyllable, the traditional meter of Bosnian/Croatian/Montenegrin/Serbian (BCMS) oral epic poetry, is even more complex, regulating multiple interactions of three prosodic dimensions, namely, stress, tone, and weight. In particular, like the Kalevala, the epic decasyllable exhibits stress-modulated weight mapping. Additionally, it exhibits a hitherto undocumented interaction type, namely, stress-modulated tone mapping. The latter constraint is further significant because it is to our knowledge the first demonstration of tone regulation in an Indo-European meter. Even though other Indo-European languages exhibit pitch accent (e.g. Ancient Greek, Sanskrit), their meters do not regulate it (Arnold, 1905; Macdonell, 1910; Maas, 1962; Raven, 1962; Allen, 1973; West, 1987; Lubotsky, 1995; Golston and Riad, 2000; Henriksson, 2022).

Aside from its contributions to metrical theory and typology, the article also bears on phonological modeling more generally, in two main respects. First, we demonstrate that violations of metrical constraints interact CUMULATIVELY, both across constraints (GANGING CUMULATIVITY) and within them (COUNTING CUMULATIVITY) (Jäger and Rosenbach, 2006; Breiss, 2020; Kawahara and Breiss, 2021; Breiss and Albright, 2022). This finding supports Harmonic Grammar (as used here) over Optimality Theory. Second, because meter is a filter superimposed on natural language, in order to determine what the meter productively contributes, we must compare the distribution of prosodic properties in meter to their baseline distribution in non-metrical language. We employ two such baselines (prose and permutation), comparing them to meter in various ways, including as priors in Maxent models, demonstrating (along with Wilson 2006; Hayes and White 2015; Henriksson 2022) the usefulness of informative priors in constraint-based modeling.

The article is organized as follows. §2 provides background on BCMS prosody and metrics. §3 presents statistical evidence for metrical regulation. In §4, we conduct additional tests to show that the observed regulation effects are productively imposed by meter, not simply reflections of ordinary-language phonology. §5 discusses the broader implications of our findings. §6 provides a formal account of the meter set in Maxent grammar. §7 concludes.

2 Background

2.1 Neoštokavian BCMS word prosody

Neoštokavian (NS), the standardized variety of BCMS, has a pitch accent system. Each prosodic word in NS has exactly one pitch-accented syllable. Two accent types are distinguished: rising and falling (Table 1; Ivić, 1958; Nikolić, 1970; Lehiste and Ivić, 1986). NS pitch accents incorporate two prosodic features: tone and stress (Inkelas and Zec, 1988; Zsiga and Zec, 2013). All pitch accented syllables bear stress, but only falling-accented syllables also feature a High tone, realized as an F0 peak on the stressed syllable. Rising-accented syllables are toneless: they exhibit a low pitch plateau in the stressed syllable, with the F0 peak on the posttonic (Lehiste and Ivić, 1986; Zsiga and Zec, 2013). Vowel length is contrastive in stressed and posttonic vowels. Codas do not contribute to weight. That is, only syllables with long vowels are heavy (Zec, 2000).

Orthography	Accent Type	Example	IPA	Gloss
â	long falling	$\boxed{\text{sûn}}$ce	[ˈsúun.tsɛ]	‘sun.NOM.SG’
ȁ	short falling	$\boxed{\text{krȁ}}$va	[ˈkrá.ʋa]	‘cow.NOM.SG’
á	long rising	$\boxed{\text{rú}}$ka	[ˈruu.ká]	‘arm.NOM.SG’
à	short rising	$\boxed{\text{vò}}$da	[ˈʋɔ.dá]	‘water.NOM.SG’
ā	long unaccented	ȉ$\boxed{\text{grā}}$	[ˈí.graa]	‘play.PRS.3SG’
ă	short unaccented	krȁ$\boxed{\text{va}}$	[ˈkrá.ʋa]	‘cow.NOM.SG’

Table 1: BCMS pitch accents.

Pitch accents are subject to strict distributional constraints (Table 2). Monosyllables can only bear falling accents. In polysyllables, accent type (rising vs. falling) is only contrastive in the initial syllable; medially, only rising accent is permitted, and finally, no accent is allowed.

	ONLY (MONOSYLLABLE)	INITIAL	MEDIAL	FINAL
FALLING	✓	✓	✗	✗
RISING	✗	✓	✓	✗

Table 2: Distributional restrictions on NS pitch accents.

Together, stress, tone, and weight define a prominence hierarchy for BCMS pitch accents (Table 3). Long falling accents are maximally prominent, being simultaneously stressed, High-toned, and heavy. On the opposite side, short rising syllables bear stress, but lack prominence in the other two categories. In between are short falling and long rising accents, which are doubly prominent.

Pitch accent	stress	High tone	vowel length
long falling	+	+	+
short falling	+	+	−
long rising	+	−	+
short rising	+	−	−

Table 3: Prominence hierarchy in BCMS.

2.2 Epic decasyllable

The epic decasyllable is the most extensively studied verse form in BCMS folk poetry. Poems in this rich oral tradition were performed live by folk poets known as *guslars*, named after the instrument that accompanied their performance—the gusle. The epic decasyllable is primarily associated with heroic (or "masculine") poetry. Still, a non-negligible number of lyric (or "feminine") poems are also composed in the meter.

The first systematic description of the epic decasyllable was offered in the foreword to the Leipzig edition of Vuk Karadžić's *Serbian Folk Poems* (Karadžić, 1824, liii-liv). Karadžić identifies central properties of the meter: Each line comprises ten syllables. The caesura (obligatory break) invariably follows the fourth syllable, dividing the line into two uneven hemistichs (3). Unmetrical lines are exceedingly rare (Foley, 1993, 88-94).

(3) $\sigma\sigma\sigma\sigma\|\sigma\sigma\sigma\sigma\sigma\sigma$

Drawing a distinction between recitation (*kazivanje*, "saying") and musical performance with the gusle (*pjevanje*, "singing"), Karadžić (1824) observes that recitation follows the prosody of ordinary language, whereas sung performance adheres to a strict metrical template, with ictus on odd-numbered syllables.[1] Accordingly, the line is a trochaic pentameter in sung performance (Karadžić, 1824; Lord, 1960). It consists of five trochaic feet, with a pause after the second foot (Figure 1) (also Hayes 1988, 224).

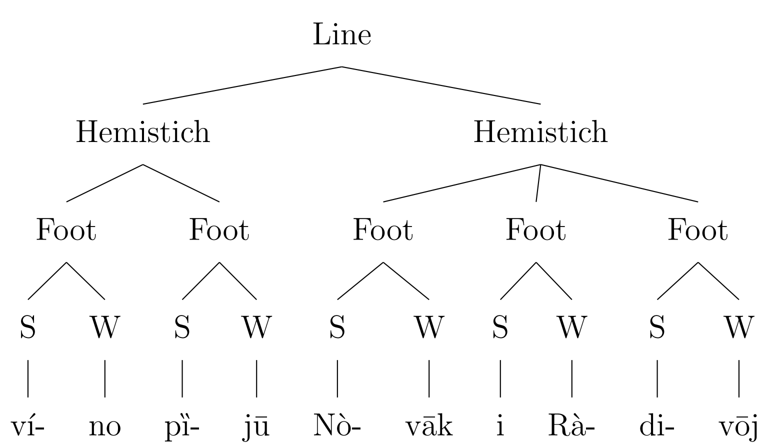


Figure 1: Epic decasyllable as trochaic pentameter.

[1]Cf. https://www.youtube.com/watch?v=jse8yzSH2WI.

Accented (i.e. stressed) syllables have been found to preferentially map to strong positions (Jakobson, 1966; Ružić, 1975; Batinić, 1975), which would place the epic decasyllable in the category of accentual/syllabotonic meters. However, the mapping of stress to the metrical template in Figure 1 is far from categorical, and inconsistent across the line. Distinguishing between metrical constants and tendencies, Jakobson (1966) rightly regards stress alignment as a metrical tendency. Mismatches between the proposed trochaic template and BCMS stress were observed as early as Karadžić (1824, liii), who reports that stressed syllables are tolerated in weak positions.[2]

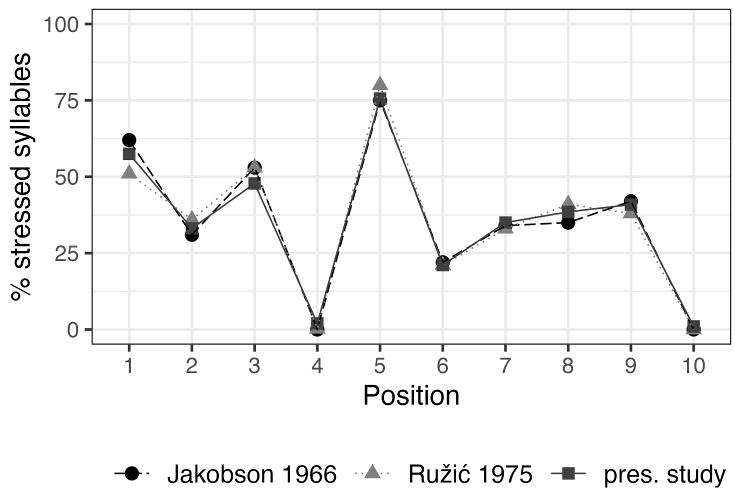


Figure 2: Proportion of stressed syllables (vertical axis) by position (horizontal axis) and source (line type).

Figure 2 plots the by-position distribution of stressed syllables in the epic decasyllable based on three sources: Jakobson (1966), Ružić (1975), and the present study. In all studies, the distribution of stressed syllables generally follows a zigzag pattern, with peaks in odd (i.e. metrically strong) positions, the sole exception being the fourth foot (positions 7 and 8). Hemistich-final positions are invariably unstressed for a meter-independent reason: They coincide with word-final syllables, which cannot bear stress in BCMS. Stressed monosyllables, for their part, are also barred from these positions (Maretić, 1907; Jakobson, 1966; Zec, 2008), a tendency observed in many of the world's meters (e.g. in Latin, Ancient Greek, Tocharian B, and Kalevala Finnish). Given the relatively high proportion of pitch-accented syllables in even positions, Zec (2008, 69) concludes that, based on stress mapping, the trochaic tendency of the meter is "fairly weak."

In the absence of categorical stress regulation, a strain of work has been skeptical of any periodic rhythmic organization in the epic decasyllable (Maretić, 1907; Vaillant, 1932; Matić, 1964; Foley, 1993). On this view, stress mapping and apparent trochaic rhythm in

[2]Strictly speaking, Karadžić (1824) discuses the metrical treatment of long vs. short syllables. However, it is clear from his examples that his "long" and "short" refer to stressed (i.e. pitch-accented) and unstressed (i.e. unaccented), respectively. His terminology is influenced by Classical metrics, where long vs. short is the operative distinction in prominence.

the meter are epiphenomenal, reflecting distributional tendencies inherent to BCMS. That is, the meter is trochaic because BCMS is overwhelmingly trochaic (Matić, 1964, 332).

Against this backdrop, several distributional patterns in the epic decasyllable provide more fine-grained evidence for stress regulation. First, the meter overrepresents "trochaic" (initially stresed: $ˈ\sigma\sigma\sigma\sigma$) quadrisyllabic prosodic words over their "iambic" (peninitially stressed: $\sigma ˈ\sigma\sigma\sigma$) counterparts (Ružić, 1975, 93). This is relevant because quadrisyllables virtually always begin in odd positions in the epic decasyllable,[3] so that "iambic" quadrisyllables induce stress misalignment. In particular, in prose, peninitially stressed quadrisyllables greatly outnumber (by 3.3 times) initially stressed ones. In meter, by contrast, the two types are approximately equally frequent (Table 4).[4]

word type	prose	epic decasyllable	prose/epic ratio
$ˈ\sigma\sigma\sigma\sigma$	3.5%	12.6%	**1:3.6**
$\sigma ˈ\sigma\sigma\sigma$	11.6%	12.4%	**1:1.07**

Table 4: Trochaic vs. iambic quadrisyllables in the meter vs. prose (Ružić, 1975).

Ružić (1975, 94) offers another test for stress regulation, namely, trisyllable sequencing in the second hemistich. When the hemistich contains two trisyllables, one initially stressed ($ˈ\sigma\sigma\sigma$) and the other peninitially stressed ($\sigma ˈ\sigma\sigma$), the unmarked order is (4a), locating both stressed syllables in odd positions. The reverse order in (4b) is marked, as it locates both stresses in even positions. In our corpus, the unmarked order (4a) occurs almost twice as often as the marked order (4b) (68 vs. 36 tokens). In our prose sample (see §4.1), by contrast, the same ratio is 15 to 13.

(4) a. $\sigma\sigma\sigma\sigma \parallel ˈ\sigma\sigma\sigma\ \sigma ˈ\sigma\sigma$
b. $\sigma\sigma\sigma\sigma \parallel \sigma ˈ\sigma\sigma\ ˈ\sigma\sigma\sigma$

Finally, the meter excludes stressed monosyllables from weak positions, except in hemistich-initial feet (Jakobson, 1966; Zec, 2008). Per Zec (2008), this constraint further substantiates the trochaism of the meter.

In conclusion, a few previously applied tests suggest that the folk poets intentionally employ word shapes and word orders in ways that induce trochaic rhythm. As this article will reinforce with further evidence, the epic decasyllable is a trochaic pentameter.

[3]Of 1,694 quadrisyllables in our corpus, only one starts in an even position.

[4]We replicated Ružić (1975)'s quadrisyllable test on our corpus and obtained a still significant, though more modest result (odds ratio 1.43, $p = 0.01$). For $\sigma ˈ\sigma\sigma\sigma$, our epic/prose ratio of 1:1.03 closely matches Ružić's of 1:1.07. However, for $ˈ\sigma\sigma\sigma\sigma$, our ratio is 1:1.4 (vs. Ružić's 1:3.6).

3 Corpus study

3.1 The corpus and its annotation

Our metrical corpus comprises 3,771 lines randomly selected from Books 1–4 of the Vienna edition of *Serbian Folk Poems* compiled by Vuk Karadžić (Karadžić, 1841–1862).[5] This corpus is substantially larger than those used in previous studies (cf. 753 lines in Jakobson, 1966; 435 in Ružić, 1975). Further, it includes poems from multiple thematic cycles.

All syllables were annotated for pitch accent (i.e. stress and tone: unaccented, rising, and falling) and vowel length (short vs. long). Coding was done using prosodic information in the second edition of Vuk Karadžić's (1852) *Serbian Dictionary*, available on the Raskovnik platform (https://raskovnik.org). Karadžić (1852) was chosen over contemporary BCMS dictionaries because it is diachronically closer to the NS dialects spoken by the folk poets whose work Karadžić documented, and because it includes many epic-specific lemmata absent from contemporary dictionaries.

Accent was annotated in two manners, namely, ISOLATION and CONTEXTUAL. Isolation accent marks stress and tone as given in the dictionary. Contextual accent marks changes in stress/tone as a function of the words' phrasal environments. In practice, the two versions differ only in cases of accent retraction to a proclitic. In the NS dialects of BCMS, word-initial falling accents may retract to a proclitic. Two types of retraction are attested. In full retraction (5a), both stress and High tone shift to the proclitic, producing a falling accent on it. In partial retraction (5b), only stress is retracted, while High tone remains on the word-initial syllable, resulting in a rising accent on the proclitic. Both patterns are optional and vary considerably across individual NS dialects, and in some cases even within a single dialect (Selkirk, 1996; Zec, 2005; Werle, 2009).

(5) a. Full retraction
   [vȍdu] 'water.ACC.SG' [ȕz=vodu] 'by=water.ACC.SG'

   b. Partial retraction
   [jȅzeru] 'lake.LOC.SG' [nà=jezeru] 'in=lake.LOC.SG'

The difference between full and partial retraction follows from underlying representation. Forms that undergo full retraction are underlyingly toneless, thereby receiving default initial stress and High tone, as in (6a). By contrast, forms displaying partial retraction have an underlying High on the initial mora (Inkelas and Zec, 1988; Zec, 2005), as in (6b).

(6) a. Toneless forms: Full retraction
   /ʋɔdu/ [ˈʋɔ́du] [ˈúz=ʋɔdu]

   b. Undelrying High: partial retraction
   /jɛ́zɛru/ [ˈjɛ́zɛru] [ˈna=jɛ́zɛru]

Our analysis avails itself of contextual accent, following Maretić (1901); Ružić (1975); Batinić (1975). That the meter utilizes contextual rather than isolation accent is evidenced

[5] We used the 1972 reprint edition. The sample included 989 lines from Book 1 (poems 340–342, 637–641; 711–713; 715–725; 727), 638 from Book 2 (poems 30–32), 1,525 from Book 3 (poems 11, 24, 45, 47, 76, 86), and 621 from Book 4 (poems 1, 3, 11, 13–15).

by stressed monosyllables. While monosyllables, as discussed, are generally prohibited in hemistich-final positions 4 and 10 (as well as in position 8), some instances can be found in those positions. In all such cases, the monosyllable is preceded by a proclitic, which receives stress in that context. Accent retraction ensures that the otherwise strict ban on stressed monosyllables in weak positions is maintained.

Following Zec (2005), we treat the verbal negation marker *ne* [nɛ] as a prefix rather than a separate word, despite the fact that *ne* is orthographically separated from the verb. Initial falling accents shift to the negation marker without exception, even in dialects that never display retraction to proclitic in cases like (5). This suggests that the verbal negation marker is different from regular prolitics (7). We therefore treat form (7b) as an initially-accented trisyllable rather than an accented monosyllable followed by an unaccented disyllable. In this manner, we avoid treating a number of apparent monosyllables as rising-accented, in clear violation of the principles of accent distribution in BCMS (§2.1).

(7)	a.	râdīm	[ˈráadiim]	‘work.PRS.1SG’
	b.	nè rādīm	[ˈnɛráadiim]	‘NEG.work.PRS.1SG’

3.2 Results

We begin our examination of the annotated metrical data using descriptive statistics conducted in R (R Core Team, 2023). Hemistich-final positions (4 and 10) are excluded from all analyses because they cannot bear stress or tone due to independent factors (§2.2).

Figure 3 shows the aggregate rates at which syllables of different types are found in odd (strong) as opposed to even (weak) positions. For example, 91% of syllables that are both heavy and falling (and therefore also stressed) occupy strong positions.

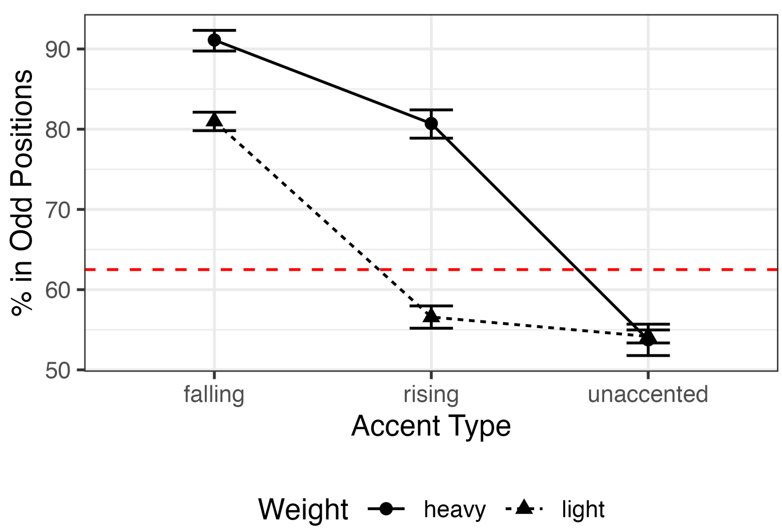


Figure 3: Proportion of syllables in odd positions (vertical axis) by accent type (horizontal axis) and weight (line type). Error bars represent Wilson 95% confidence intervals. The dashed rule marks the chance level (grand mean).

All rates in Figure 3 are above 50% because the hemistich-final positions, both even/weak, are excluded: There remain five odd (strong) positions and three even (weak) positions. If syllable types were randomly distributed, they would cluster around $5 \div 8 = 0.625$ (dashed horizontal bar). Figure 3 suggests that weight plays a role among stressed (i.e. falling and rising) syllables, but not among unstressed (unaccented) syllables. Further, tone appears to be relevant, such that within each weight class, falling is more strong-biased than rising. Indeed, rising lights are *below* the chance baseline.

We assess these results statistically using mixed-effects logistic regression (`lme4`; Bates et al. 2015), which permits further controls. The dependent variable is metrical strength, coded as "1" for odd/strong and "0" for even/weak. The fixed effects include syllable weight, accent type, and their interaction. Weight is coded as "0" for light and "1" for heavy (i.e. long-voweled). Accent type has three levels, namely, unaccented, rising (stressed toneless), and falling (stressed High), forward-difference coded such that rising is compared to unaccented, and falling to rising (as opposed to all levels against the baseline level of unaccented, as would be the case with dummy coding). Tone is not modeled separately because doing so would violate independence: Because High is possible only in stressed or posttonic position, whenever a stressed toneless syllable occupies a strong position, an unstressed High occupies the following weak position, and vice versa.

The model also includes random intercepts for WORD SHAPE, encoding three properties of each syllable datum: (i) its position in the word, (ii) its word's size, and (iii) its word's weight template (excluding the target syllable, whose weight is modeled as a fixed effect). Given these specifications, word shape has 68 unique levels (e.g. X-light, light-X-heavy, X, etc.).[6] The model formula and results are provided in Table 5. The high standard deviation of the random intercepts suggests that the effects of accent type and syllable weight on metrical position indeed vary considerably across word shapes.

Random effects	*Variance*	*Standard Deviation*		
Word shape (N = 68)				
(Intercept)	48.25	6.95		
Fixed effects	β	*Standard Error*	*Wald z*	*p*
(Intercept)	0.72	0.79	0.91	0.36
Accent type (difference coded)				
Rising (vs. Unaccented)	0.26	0.06	4.34	0.000 ***
Falling (vs. Rising)	0.98	0.07	14.77	0.000 ***
Weight (baseline: Light)				
Heavy	0.26	0.05	5.13	0.000 ***
Accent type*Weight interaction				
Rising (vs. Unaccented):Heavy	0.69	0.11	6.15	0.000 ***
Falling (vs. Rising):Heavy	−0.77	0.14	−5.57	0.000 ***

Table 5: Mixed-effects logistic regression model output. Model formula: `strong ~ accent_type * weight + (1 | shape)`.

[6]Because the model takes each syllable as a datum, it would not make sense to use full words (e.g. *bukovu*, *tiho*, *besjedila*) as random effects. Each word would effectively be assigned a propensity to appear in strong vs. weak positions, which makes little sense for polysyllables, which span multiple positions.

Turning to the main effects, tone is significant in stressed syllables: Falling (stress + High) is more strong-aligned than rising (stress + no tone); rising, in turn, is more strong-aligned than unaccented. Weight is also significant, such that heavies are more strong-aligned than lights. Additionally, the two interactions between accent type and weight are significant. The first, being positive, indicates that the effect of weight is stronger in rising syllables than unaccented syllables. The second, being negative, indicates that the effect of syllable weight is attenuated in falling compared to rising syllables.

The effect of weight within each accent type taken separately is assessed by Tukey *post hoc* pairwise comparisons. In unaccented syllables, heavies do not differ significantly from lights ($\beta = -0.06, p = 0.92$). In rising syllables, heavies are significantly more strong-aligned than lights ($\beta = -0.75, p < 0.0001$). Finally, in falling syllables, there is no significant effect of weight ($\beta = 0.02, p = 1$).[7]

Taking stock, roles for both accent type and weight are supported for meter. Moreover, the two factors interact, such that the effect of weight is attenuated closer to floor and ceiling levels, producing the characteristic sigmoid sometimes described as a "wug shape" (Zuraw and Hayes, 2017; Kawahara, 2020; Hayes, 2022).

4 Further evidence for regulation: baselines of comparison

4.1 Confounding factors

The corpus analysis in §3.2 found effects of both accent type and weight in the BCMS epic decasyllable, confirming that not just stress, but also tone and weight should be taken into account (Batinić, 1975). The model included random intercepts for word shape, reducing its sensitivity to specific word shapes' idiosyncratic distributions. In this section, we control for potential confounds more directly using baselines of comparison (i.e. observed vs. expected studies), further supporting the same findings.

The issue is that independent of prominence mapping, the epic decasyllable imposes constraints on the distribution of word shapes (Maretić, 1907; Jakobson, 1966; Foley, 1993; Zec, 2008). To give just a couple of examples, in §2.2, it was mentioned that monosyllables virtually never occur hemistich-finally, and quadrisyllables virtually never start in even positions. (More generally, the longer the word, the more likely it is to be localized line-finally.) What's more, BCMS pitch accents are subject to strict distributional constraints. For example, falling accents are restricted to word-initial position (§2.1). The observed positional contrasts between falling vs. rising accents might therefore be epiphenomenal, reflecting the skewed distributions of word shapes and the skewed distributions of accent types within word shapes. To give a simple example, because monosyllables can only bear falling accent, falling accents will be more frequent in any position in which monosyllables are more frequent, regardless of the cause.

[7] While this lack of a contrast between heavies vs. lights in falling syllables might appear to be inconsistent with the visual gap between them in Figure 3, recall that the figure plots the raw empirical distribution, while the model corrects for word shape.

To control for the distributional restrictions on BCMS pitch accents, we examine a restricted set of environments, where both falling and rising accents are available. Further, to gain a better understanding of the baseline prosodic distributions in the language, we compare our metrical corpus to control samples of non-metered BCMS texts.

4.2 Control samples

Comparison with non-metered texts allows us to test whether the poets deliberately place prominent syllables in odd positions, or whether such syllables are naturally inclined towards odd positions for reasons independent of the meter. We use two control samples: a prose selection and a scrambled (randomly permuted) version of the metrical corpus. Both methods have a long-standing tradition in metrics (see Tarlinskaja and Teterina, 1974; Tarlinskaja, 1976; Bailey, 1975; Gasparov, 1980, 1987; Hall, 2006; Hayes and Moore-Cantwell, 2011; Hayes and Schuh, 2019; Henriksson, 2022 for prose comparison and Janson, 1975; Gunkel and Ryan, 2011; Ryan, 2017 for scrambled metrical corpora).

First, we use the so-called "Russian method," comparing metered verse to poetry-like sentences from prose texts (Bailey, 1975; Tarlinskaja, 1976; Gasparov, 1980; Hayes and Schuh, 2019). We extracted epic decasyllable-like sentences from a diverse sample of BCMS novels and short stories (late 19th to early 21st centuries). Specifically, we retrieved 801 prose sentences of exactly ten syllables with a word boundary after the fourth syllable, matching the meter's caesura. This practice follows Gasparov (1987)'s suggestion that prose comparanda should comply with the rules of the meter that are independent of those being tested. The sentences were prosodically annotated as in §3.1.

The scrambled metrical corpus was constructed following the "Rigged Veda" method of Gunkel and Ryan (2011). For each real line in the epic corpus, a program generates a new line by replacing each real word with a word of the same shape (i.e. number of syllables, stress position, and weight template) randomly selected from the corpus. Ten lines were generated for each real line, for a total of 37,710 new lines. The scrambled corpus preserves the original's distribution of word boundaries, stressed positions, and syllable weight, but not tone (since tone is what will be tested). For example, for the real line template (8) (e.g. I, 717:1), the program generates permuted lines such as those in (9).

(8) fu rU ‖ u fuu Ru (f = falling, r = rising, u = unaccented, capitalization = heavy)

(9) a. fu fU ‖ u fuu Fu
b. fu fU ‖ u fuu Ru
c. ru fU ‖ u ruu Fu

4.2.1 Prose comparison: overall results

We first turn to the aggregate effects of stress, tone, and weight in the epic vs. prose (Figure 4). (The epic panel repeats Figure 3 above.) As in §3.2, hemistich-final positions are excluded.

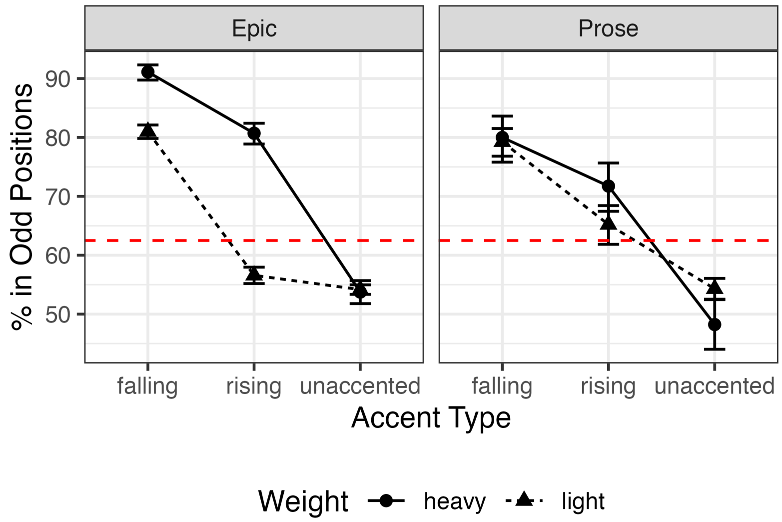


Figure 4: Proportion of syllables that occur in odd positions (vertical axis), by accent type (horizontal axis) and syllable weight (line type) in the oral epic vs. the epic-like prose sample (facets). The dashed rule indicates chance level.

First, consider weight. In the epic, there is a clear effect of weight in stressed (i.e. falling or rising) syllables: Heavies (solid line) are well separated from lights (dashed line). In prose, by contrast, hardly any separation between heavies and lights is apparent for any accent type.[8] This confirms that the differential treatment of heavies vs. lights in specifically stressed syllables in the epic is not merely a reflex of the language more generally.

Next, consider tone. On the one hand, the comparison with prose underscores the importance of undertaking such a comparison in the first place: If one were to observe only the epic, one might conclude that there is strong evidence for tonal sensitivity in both heavy and light syllables. But the prose distribution reveals that at least some of that apparent contrast is epiphenomal. On the other hand, there is still a notable difference in the distribution of tone between the two corpora. The dashed line for lights has an elbow in the epic that is not mirrored by the prose, suggesting that falling vs. rising is actively regulated by the meter. More specifically, rising lights are below the grand mean in the epic (56.6%) but above it in prose (65.2%) (Fisher's exact test OR $= 0.70, p < 0.0001$). Regression likewise bears this out. When the mixed model used in §3.2 is run on the prose, falling vs. rising is non-significant ($\beta = 0.18, p = 0.15$). The following sections will corroborate the conclusion that tone is actively regulated by the meter.

Since stress does not appear to be regulated on its own, it is further worth testing whether any stress is needed in the model at all, i.e. whether all observed patterns can be explained solely by weight and tone. To test this, we fit regressions with and without stress as a predictor. The former is simply the model already presented in §3.2. The stressless model includes fixed predictors for tone (High vs. none, where both falling and post-rising syllables

[8] More precisely, when the regression model from §3.2 is run on the prose sample, weight is weakly significant as a main effect ($\beta = 0.21, p = 0.02$), but more importantly, neither of the interactions with accent type and weight is significant ($\beta = 0.06, p = 0.79$; $\beta = -0.3, p = 0.20$).

were coded as High), weight, and their interaction, as well as with random intercepts for word shape, as before. The stressless model substantially underperforms the model with stress, the BIC (Schwarz, 1978) of the former being higher by 317.

To conclude, prose comparison supports the regulation of weight in stressed syllables. Stressed heavies display a robust affinity towards strong positions in the meter, exceeding their prose rates. Evidence for tone regulation is suggestive but less conclusive at this point, judging by the visualizations, but we present more rigorous evidence supporting it in §4.2.3.

4.2.2 Stress-weight interaction: by-position analysis

As established, stressed heavies are aggregately overrepresented in strong positions in the epic. This pattern is not matched by prose, where stressed heavies and lights appear in strong positions at roughly equal rates (Figure 4). We next analyze the distribution of stressed heavies across individual positions of the line. After all, the aggregate effect might be driven by only certain positions. In particular, Jakobson (1966, 418–419) posits a "quantitative clausula": Positions 7 and 8 tend to be light when stressed, while position 9 (in the line-final foot) tends to be heavy when stressed (see also Maretić 1901; Zec 2008).

Figure 5 shows the rate at which stressed syllables are heavy across positions, both in the epic (top) and in prose (bottom). The results support Jakobson (1966)'s observation. Positions 7 and 8 display the lowest rates of stressed heavies. By striking contrast, in position 9, 84% of stressed syllables are heavy, triple or quadruple the rate anywhere else in the line. This stringency of the final foot is a manifestation of the cross-linguistic principle known as FINAL STRICTNESS, whereby meters tend to be (if anything) stricter at the ends of lines (Arnold, 1905; Kiparsky, 1968; Hayes, 1983; Ryan, 2017, 2019; deCastro Arrazola, 2018; Smith and Anttila, 2025).

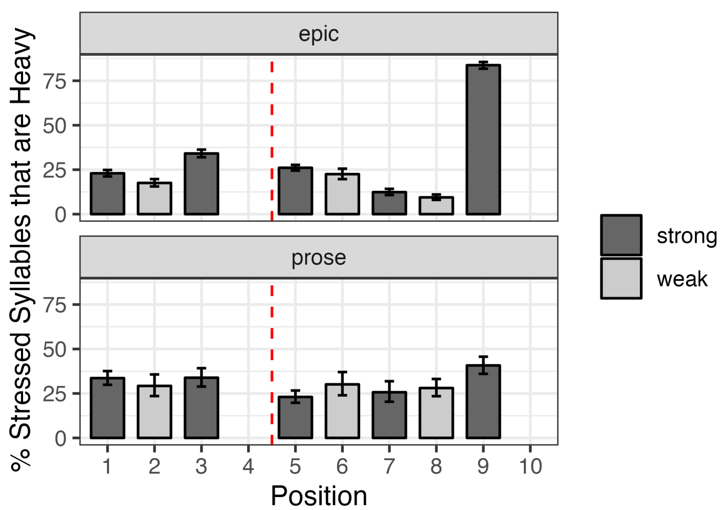


Figure 5: The proportion of stressed syllables that are heavy (vertical axis) by position (horizontal axis) in epic vs. prose (facets). The dashed vertical indicates caesura.

The strong effect at the end of the line (as previously recognized) raises the question of whether weight is regulated at all in stressed syllables earlier in the line (not previously recognized). We find that weight is indeed regulated (albeit more weakly) even outside of the cadence (final foot). For starters (see also the Maxent models in §6), this is evident from Figure 5, where within every foot of the line, the stressed heavy rate is higher in the strong position than in the following weak position (if data are available). This consistent directionality contrasts with prose, which shows no systematic odd–even asymmetry.

Figure 6 repeats the information from Figure 5, but shows the epic-to-prose ratio directly. Compared to prose, stressed heavies in the epic are underrepresented in all weak positions, as well as in strong positions 1 and 7. Positions 3 and 5 are roughly at the prose baseline. Position 9 stands out, with a substantial overrepresentation of stressed heavies in the epic.

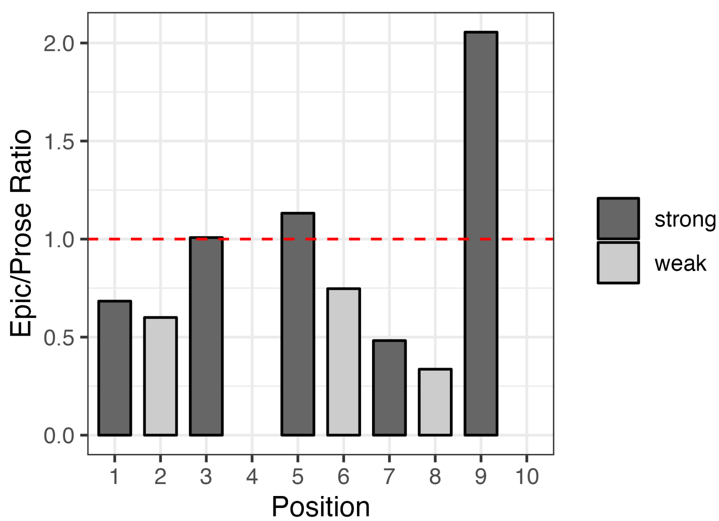


Figure 6: Epic/prose ratio of stressed heavy syllables (vertical axis) by position (horizontal axis). The dashed horizontal rule is the baseline.

We assess the patterning of stressed heavy syllables in the precadence (positions 1–8) using a mixed-effects logistic regression. As before, the model excludes positions 4 and 10, but now the model also excludes the entire cadence (i.e. position 9 as well). The model is otherwise set up exactly as in §3.2. In this new model, the main effect of weight disappears: $\beta = -0.03, p = 0.60$. However, the interactions between weight and accent type persist, with the same directionalities: rising (vs. unaccented) * heavy ($\beta = 0.58, p = 0.000$) and falling (vs. rising) * heavy ($\beta = -0.75, p = 0.000$), confirming that stress-weight interactions are not confined to the cadence.

It remains to be explained why stressed heavies are underrepresented in positions 1 and 7, despite their being strong. We propose that the general scarcity of stressed heavies in the precadence reflects what Ryan (2017) calls "making good use of the available lexicon." In general, the epic poets utilize heavy-stressed words at roughly the same rate as do prose authors: They comprise 29% of words in the epic vs. 32% in prose ($\text{OR} = 0.94, p = 0.18$). In the epic, because heavy-stressed disyllables are strongly biased towards the cadence, they are less likely to appear earlier in the line. This depletion primarily affects positions nearby the cadence (e.g. 7 as opposed to 5) because the reordering of words within phrases is one way in which it is achieved. For example, if the line-final phrase is a heavy-light adjective modifying a light-light noun, the marked noun-adjective order is more likely to be employed.[9]

In summary, weight is regulated most stringently in the cadence, but not exclusively there. Stress-weight interactions manifest in two ways. First, stressed syllables assigned to the head of the final foot tend strongly to be heavy. Second, stressed heavies are avoided in weak positions throughout the line. In Figure 6, for instance, the dark bars are closer to the baseline than their light counterparts.

4.2.3 Stress-tone interaction

Having treated weight in the previous sections, we now turn to tone. In BCMS, as established, the only environment in which accent type (falling vs. rising) is contrastive is in the initial syllables of polysyllables. We therefore begin by examining the initial syllables of specific word shapes, controlling for both word length (disyllable or trisyllable) and the weight of the initial syllable.

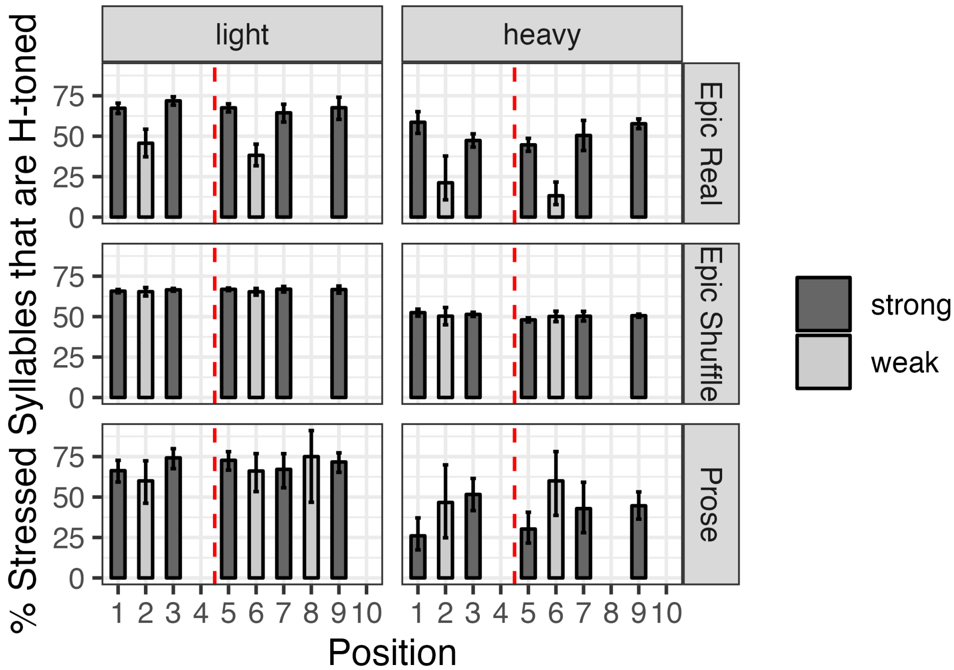
Figure 7 depicts the rate at which stressed initial syllables bear falling vs. rising tone across positions of the line. The real epic rates are shown in the top row, followed by the rates from two comparison corpora, namely, the word-order permuted epic (epic shuffle) and prose. Word shapes are considered separately to eliminate any possible confounds from the different distributions of different word shapes (Ryan, 2011). Disyllables are in the top half of the figure, followed by trisyllables. The weight of the initial syllable is also controlled by separating light (left half) from heavy (right half).

The real epic corpus exhibits a zigzag pattern. The proportion of High-toned stressed syllables peaks in strong positions and dips in weak ones. This generalization holds in three of the four quadrants of Figure 7. (No effect is apparent for heavy-initial trisyllables, for which data are sparse; note the large error bars.) By contrast, the control corpora show no consistent effect of position type. In the shuffle, the rate is relatively flat across positions. In the prose, the rate is more volatile, with no consistent trend.

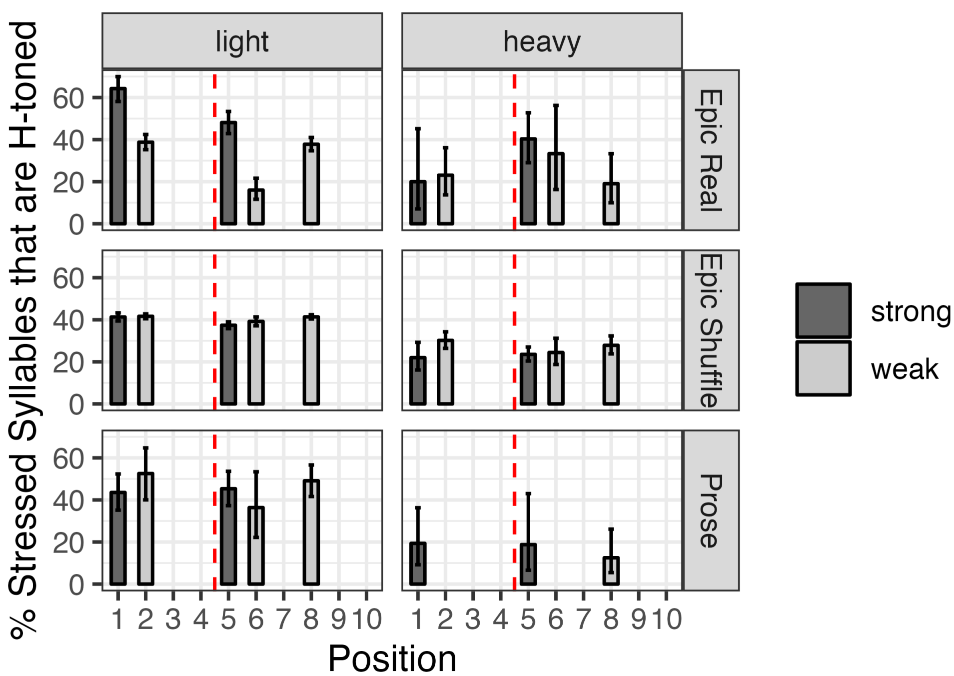
Thus, controlling for the distributional constraints on pitch accent, an effect of tone emerges: High-toned stressed disyllable and trisyllable initials are biased towards strong over weak positions. Particularly instructive is that disyllable and trisyllable initials pattern alike with respect to tone, despite having nearly inverse distributions: Disyllable initials overwhelmingly occupy odd positions (91%), vs. 23% for trisyllable initials. This consistent treatment of tone across largely independent contexts provides strong evidence that tone regulation is independent of word shape/boundary distribution.

Tone appears to be the most stringently regulated in weak positions. That is, falling syllables are more disfavored in weak positions than rising syllables are in strong positions. In Figure 7, this is evident from the fact that the real rates in strong positions (dark bars) are at roughly the same levels as the control corpora, whereas the rates in weak positions (light bars) are markedly lower in the real corpus. (The Maxent model below (§6.1.2) will reinforce this asymmetry statistically.)

[9] The fact that position 1 is significantly lower than position 3 might likewise reflect a weaker effect of constituent-final strictness effect within the first hemistich.



(a) Disyllable initials



(b) Trisyllable initials

Figure 7: Proportion of High-toned vs. toneless stressed initial syllables (vertical axis) by Position (horizontal axis), faceted by syllable weight (columns) and corpus (rows). The dashed verticals indicate caesura.

4.3 Section conclusion

The distribution of weight, tone, and stress in the epic meter is influenced not only by prosodic preferences specific to the meter, but also by the inherent distribution of these features in BCMS prosody along with independent distributional constraints on word shapes (e.g. a break must appear after position 4, a hemistich must not end with a monosyllable, longer words tend to biased towards the end of the line, etc.). To isolate the meter's contribution to the distribution of prosodic features, we conducted a series of controlled tests, using two corpora as baselines, namely, a scrambled version of the corpus and a sample of epic decasyllable-like sentences extracted from prose.

The inferential statistics support the following two main generalizations. First, the meter is sensitive to weight in stressed syllables. While stressed syllables show a particularly strong tendency to be heavy in position 9 (the final strong), weight is also regulated elsewhere in the line. Second, stressed High-toned (i.e. falling) syllables are avoided in weak positions. Both generalizations are evident from comparisons to the baselines (this section), and will be further reinforced by the constraint selection procedure in §6.1.2. BCMS folk meter, we conclude, regulates doubly-prominent syllables, that is, syllables bearing both stress and weight, or both stress and High tone.

5 Discussion

The epic decasyllable has proven difficult to situate in metrical typology (Jakobson, 1966; Zec, 2008; Kiparsky, 2020). This difficulty stems primarily from the complexity of BCMS word prosody. The language features stress, tone and vowel length, which interact to form a prominence hierarchy (§2.2). The sheer number of prosodic features and their combinations in BCMS makes it challenging to determine which of them, if any, are regulated in the meter. There has thus been little consensus on regulation effects in the epic decasyllable. Our findings contribute to this debate by providing evidence that the meter regulates both weight and tone specifically in stressed syllables.

5.1 Metrical typology

Typological classification of meters is based primarily on the parameters in (10) (Hanson and Kiparsky, 1996).

(10) a. PROMINENCE TYPE: which property of syllables does the meter regulate?
b. REGULATION SITE: does the meter impose restrictions on strong or weak positions (or both)?

In terms of (10a), perhaps most meters that regulate properties of syllables such as stress, tone or weight impose restrictions on only one of these features. Depending on which prosodic property they regulate, meters are classified as accentual, quantitative, or tonal (Hanson and Kiparsky, 1996; Fabb and Halle, 2008; Blumenfeld, 2016; Kiparsky, 2020). In contrast to single-feature mapping, hybrid meters impose constraints on multiple prosodic dimensions (e.g. both stress and weight). Hybrid meters fall into two subtypes: those that exhibit independent mapping of each feature, and those with interactive mapping, where the meter regulates combinations of prosodic features (Ryan, 2017).

In terms of (10b), meters either require that strong positions be associated with prominence (stress, heaviness, or High tone), or avoid prominence in weak positions. Homeric Greek dactylic hexameter exemplifies the regulation of strong positions: each strong position must be filled by a heavy syllable (11).

(11) STRONG⇒HEAVY
If a position is strong, it must contain a heavy syllable.

English iambic pentameter primarily regulates weak positions. In this accentual meter, stressed syllables are avoided in weak positions, but there is no requirement that strong positions contain a stressed syllable (§12; e.g. Blumenfeld, 2016, cf. Hayes et al., 2012).

(12) STRESS⇒STRONG
If a syllable is stressed, it must be assigned to a strong position.

Constraint (11) penalizes strong positions that lack prominence (specifically, weight). By contrast, (12) penalizes prominent syllables in weak positions.

5.2 The place of the epic decasyllable in metrical typology

The BCMS epic decasyllable is a hybrid meter, simultaneously regulating stress, weight and tone. The mapping is interactive rather than independent, because the meter strictly regulates syllables that are concurrently prominent in at least two categories: stressed heavies and stressed High-toned syllables. In other words, weight and tone are regulated in stressed syllables, exhibiting what Ryan (2017) calls STRESS-MODULATED REGULATION.

Previously documented cases of interactive mapping in the meter, including Tamil and Finnish, involve stress-modulated weight regulation (*ibid.*). The strictness of weight mapping is determined by stress level. A similar pattern is observed in the BCMS epic decasyllable, where stressed heavies are avoided in weak positions (§4.2.2).

A typologically unique property of the epic decasyllable is its doubly interactive nature: stress interacts with both weight and tone. This dual interaction is significant because in the previously documented cases of interactive mapping, stress was found to only interact with weight. The identification of stress-modulated tone mapping in BCMS thus expands the typology of hybrid meters.

To our knowledge, tone-sensitivity has no known parallels in Indo-European metrical traditions. It is customarily argued that when Indo-European languages possess tone (i.e. pitch accent), as in Ancient Greek and Sanskrit, the meter shows no sensitivity to it (Arnold, 1905; Macdonell, 1910; Maas, 1962; Raven, 1962; Allen, 1973; West, 1987; Lubotsky, 1995; Golston and Riad, 2000; Henriksson, 2022). The BCMS epic decasyllable therefore constitutes the first documented case of tone regulation in Indo-European.

Moving to (10b) (the site of regulation), Zec (2008, 75-76) and Jakobson (1966, 420) observe that the epic decasyllable excludes prominence from weak positions rather than requiring it in strong positions. Our findings agree with this view. Both stressed heavies and stressed High-toned syllables are underrepresented in weak positions. At the same time, strong positions that lack prominence are not particularly rare (see e.g. Figure 2). This suggests that weak positions are the primary locus of regulation. The strong position is, however, clearly regulated in the cadence. If the head of the final foot contains a stressed syllable, it tends strongly to be heavy (§4.2.2), indicating that prominence (in this case, coincident prominence in two dimensions) is required in a strong position.

In summary, our findings substantiate the claims to the effect that the epic decasyllable is an interactive hybrid meter (Jakobson, 1966; Batinić, 1975; Hayes, 1988; Ryan, 2017; Kiparsky, 2020). Both weight and tone mapping are modulated by stress. Contributions to metrical typology include showing that a hybrid meter can display multiple interactions in parallel, and that tone (like weight) can participate in interactive (e.g. stress-modulated) mapping. Further, the meter furnishes a case of both strong and weak positions being regulated, but the former only in the cadence.

6 Constraint-based analysis

6.1 Maxent metrics

We now turn to the formal analysis of the epic decasyllable, which is couched in Maximum Entropy Harmonic Grammar (Maxent; Goldwater and Johnson, 2003; Hayes and Wilson, 2008), a probabilistic version of Harmonic Grammar (HG; Legendre et al., 1990). Our Maxent analysis was conducted in R, using the `maxent.ot` package (Mayer et al., 2024).

In HG, constraints are weighted rather than strictly ranked as in Optimality Theory (Prince and Smolensky, 2004). A candidate's Harmony score H_i is defined as the weighted sum of its constraint violations (13), where w_k is the weight of constraint k, c_{ik} is the number of violations candidate i incurs on constraint k, and C is the constraint set. The winner is the candidate with the highest Harmony score (where Harmony scores are nonpositive).

(13) $$H_i = \sum_{k \in C} w_k c_{ik}$$

Like classical HG, Maxent evaluates candidates by summing their violations scaled by constraint weights. However, unlike classical HG, Maxent does not select a single winner, but assigns probabilities to candidates based on their Harmony scores. To ensure that the predicted probabilities in a tableau sum to 1, Harmony scores are exponentiated, and the candidate probability p_i is then calculated as its exponentiated Harmony e^{H_i} divided by the sum of exponentiated Harmonies of all candidates j in the candidate set X (14).

(14) $$p_i = \frac{e^{H_i}}{\sum_{j \in X} e^{H_j}}$$

Maxent is well suited to modeling variation and gradience in phonology, which is precisely what our data warrant. Maxent has been argued to be superior to other frameworks used to model variation in phonology (Zuraw and Hayes, 2017; Smith and Pater, 2020; Flemming, 2021; Hayes, 2022) and is the leading approach to analyzing variation in generative metrics (Hayes and Moore-Cantwell, 2011; Hayes et al., 2012; Ryan, 2017; McPherson and Ryan, 2018; Hayes and Schuh, 2019; Henriksson, 2022; Hayes and Minkova, 2023).

In Maxent, lines' Harmony scores are a direct measure of their metricality. Metricality (i.e. Harmony) and probability are directly correlated: The more harmonic a line type, the more frequently poets are predicted to produce it. This link between metrical complexity and production frequency has been central to generative metrics since its outset, cf. the frequency hypothesis of Halle and Keyser (1971), which Maxent not only adopts but also formalizes with mathematical rigor (Hayes et al., 2012; Hayes and Schuh, 2019; Hayes, 2023).

The dominant approach in Maxent metrics (Hayes and Moore-Cantwell, 2011; Hayes et al., 2012; Ryan, 2017; Hayes and Minkova, 2023) follows the gist of Hayes and Wilson (2008)'s phonotactic models. The grammar assigns weights to constraints so as to best match the observed frequencies of line types. The model is effectively inputless. All candidate line types share a uniform dummy input and compete in a single tableau.

6.1.1 Constraints

We first define the space of potentially relevant constraints for our analysis of the epic decasyllable. Not all of them will enter the Maxent model; a selection procedure will determine which of these are justified. In defining our constraints, we adopt the format of Hanson and Kiparsky (1996); Ryan (2017). Accordingly, our constraints: (i) are structured as implications, not prohibitions, (ii) reference prominent categories (strong positions, stress, heaviness, High tone), not absence of prominence.[10]

In (15–16), we present constraints that govern the independent mapping of stress, weight and tone.

(15) a. STRESS⇒STRONG
If a syllable is stressed, it must be assigned to a strong position.

b. HEAVY⇒STRONG
If a syllable is heavy, it must be assigned to a strong position.

c. HIGH⇒STRONG
If a syllable is High-toned, it must be assigned to a strong position.

(16) a. STRONG⇒STRESS
If a syllable is assigned to a strong position, it must be stressed.

b. STRONG⇒HEAVY
If a syllable is assigned to a strong position, it must be heavy.

c. STRONG⇒HIGH
If a syllable is assigned to a strong position, it must be High-toned.

To model stress–weight interaction, we adopt the constraints in (17) (Ryan, 2017).

(17) a. STRESS⇒(STRONG⇒HEAVY)
If a syllable is stressed and assigned to a strong position, it must be heavy.

[10] These constraints could equivalently be expressed as conjunctions, since $A \rightarrow B \equiv \neg(A \wedge \neg B)$. For example, constraint (15a) could be expressed as *STRESS/WEAK, and (16a) could be expressed *UNSTRESSED/STRONG. The metrics literature (*ibid.*) has generally favored the implication format, and Ryan (2017) puts forth a substantive reason for it (namely, the restriction that mappings expressed as implications apparently never need to invoke weak elements such as light, unstressed, etc.).

b. Stress⇒(Heavy⇒Strong)
If a syllable is stressed and heavy, it must be assigned to a strong position.

Constraint (17a) penalizes stressed lights in strong positions, while (17b) penalizes stressed heavies in weak positions (18).

(18)

		Stress⇒(Strong⇒Heavy)	Stress⇒(Heavy⇒Strong)
a.	$^{\text{ˈ}}\sigma_{\mu}$/S	*	
b.	$^{\text{ˈ}}\sigma_{\mu\mu}$/W		*

The tonal counterparts of the constraints in (17) are outlined in (19).

(19) a. Stress⇒(Strong⇒High)
If a syllable is stressed and assigned to a strong position, it must be High-toned.

b. Stress⇒(High⇒Strong)
If a syllable is stressed and High-toned, it must be assigned to a strong position

Constraint (19a) militates against stressed toneless syllables in strong positions, while (19b) assigns violations for stressed High-toned syllables in weak positions (20).

(20)

		Stress⇒(Strong⇒High)	Stress⇒(High⇒Strong)
a.	$^{\text{ˈ}}\sigma_{\mu}$/S	*	
b.	$^{\text{ˈ}}\acute{\sigma}_{\mu}$/W		*

This constraint system readily captures the gradient strictness of regulation of syllable types (Figure 3). The severity of misalignment (here, placing a syllable in a weak position) correlates with its degree of prominence. BCMS syllables can bear three types of prominence: stress, weight, and tone. These dimensions combine to form a cline of markedness (§2.1): The more prominence a syllable possesses, the more costly its misalignment becomes. Stressed High-toned heavies occupy weak positions at a relatively low rate (only 9%, 167/1,880 tokens). By contrast, stressed toneless lights, with just one degree of prominence, are the stressed syllable type the most likely to be misaligned (43% misaligned); see (21).

(21)

		Stress⇒S	Heavy⇒S	High⇒S	Stress⇒(Heavy⇒S)	Stress⇒(High⇒S)	*misalignment rate*
a.	$^{\text{ˈ}}\acute{\sigma}_{\mu\mu}$/W	*	*	*	*	*	9% (167/1,880)
b.	$^{\text{ˈ}}\sigma_{\mu\mu}$/W	*	*		*		19% (371/1,923)
c.	$^{\text{ˈ}}\acute{\sigma}_{\mu}$/W	*		*		*	19% (851/4,478)
d.	$^{\text{ˈ}}\sigma_{\mu}$/W	*					43% (2,219/4,881)
e.	$\sigma_{\mu\mu}$/W		*				46% (1,115/2,497)

Finally, we include three constraints that apply only to the cadence (22), following the demonstration that weight mapping is stricter at the end of lines (Maretić, 1901; Jakobson, 1966; Zec, 2008). Constraint (22a) requires that stressed syllables in the ninth position be heavy (§4.2.2). Furthermore, constraints (22b–22c) are the cadential counterparts of (15b–16b): (22b) penalizes heavy syllables in the tenth position, while (22c) penalizes the ninth positions that do not contain a heavy syllable.

(22) a. STRESS⇒(STRONG⇒HEAVY)$_{\text{cadence}}$
If a stressed syllable is assigned to a strong position in the cadence, it must be heavy.

b. HEAVY⇒STRONG$_{\text{cadence}}$
A heavy syllable in the cadence must be assigned to a strong position.

c. STRONG⇒HEAVY$_{\text{cadence}}$
If a syllable is assigned to a strong position in the cadence, it must be heavy.

In sum, there are thirteen constraints in the initial set (15–22), which will next be subjected to a selection procedure.

6.1.2 Model selection

Using a fixed random seed for reproducibility, we randomly split the same metrical corpus of 3,771 prosodically-annotated lines used for descriptive analysis in §3–4 into two halves: a training set and a test set (n = 1,885 and 1,886, respectively). This is a HOLD-OUT, i.e. SIMPLE VALIDATION method (Devroye and Wagner, 1979; Arlot and Celisse, 2010). Constraint violations were coded automatically in R. The training set is used to select the final model and estimate constraint weights. In this subsection, we report the selection outcome and the model's fit to the training data. In §6.1.3 we assess generalization, i.e. how well the grammar predicts unseen lines, by predicting line type probabilities in the test data using the weights learned in training. Crucially, test lines do not influence constraint selection or weight estimation, which prevents information leakage and provides an unbiased estimate of predictive performance.

We determine the best-fitting grammar using stepwise-forward selection (Della Pietra et al., 1997; Hastie et al., 2009; see also Hayes et al. 2012). This is a bottom-up approach: The grammar search begins with the intercept-only model, to which constraints are added one at a time until no further improvement is possible. At each step, a single constraint is selected from the pool based on information gain (Della Pietra et al., 1997). Information gain was assessed using the likelihood ratio test (LRT). Inclusion threshold was $p < 0.05$ in the LRT. The search terminated once no remaining constraint passed the threshold.

Only five of the original thirteen constraints from §6.1.1 are selected in the final model. Table 6 reports selection history and constraint weights.

	Constraint selected	LRT	weight
Step 1	HEAVY⇒STRONG	$p < 0.0001$	0.86
Step 2	STRESS⇒(HIGH⇒STRONG)	$p < 0.0001$	1.40
Step 3	STRESS⇒(STRONG⇒HEAVY)$_{\text{cadence}}$	$p < 0.0001$	2.62
Step 4	HEAVY⇒STRONG$_{\text{cadence}}$	$p < 0.0001$	0.64
Step 5	STRESS⇒(HEAVY⇒STRONG)	$p = 0.03$	1.12

Table 6: Model selection output and history, and constraint weights.

The model selection output is fully consistent with our empirical results. Constraints on strong positions (16) were not included in the final model, confirming that weak positions are the primary locus of regulation in the epic decasyllable (§5.2). Further, both STRESS⇒(HEAVY⇒STRONG) and STRESS⇒(HIGH⇒STRONG) were selected, in line with the observed avoidance of doubly-prominent syllables in weak positions.[11] Finally, the inclusion and relative strength of STRESS⇒(STRONG⇒HEAVY)$_{\text{cadence}}$ confirms that stressed lights are robustly avoided in the ninth position (Jakobson, 1966).

The final model achieves a solid fit to the training data despite using only five constraints. Figure 8 plots observed line type probabilities in the training data against those predicted by the model.

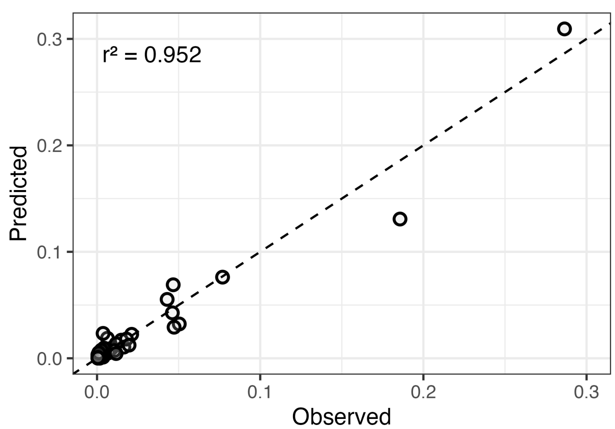


Figure 8: Scatterplot of observed vs. predicted frequencies of line types in the training dataset.

We further analyze line type frequency in the epic decasyllable based on their violation profiles. As discussed in §6.1, Maxent establishes a direct correlation between metrical well-formedness (i.e. Harmony) and production frequency of line types. Table 7 presents a selection of line types with their violation profiles.[12] The results illustrate that Harmony and observed frequency are well correlated. Violation-free candidate (a) is the most frequent line type in the corpus. For subsequent candidates, violation tallies increase and frequencies decline. Lines (b–c), which have relatively low complexity (i.e. markedness; Halle and Keyser 1971), still have triple-digit frequency. Mid-complexity lines (roughly d–f) are less frequent, and the

[11] While the p-value of STRESS⇒(HEAVY⇒STRONG) is borderline here, see §6.2 for further support.
[12] Line type frequencies in Table 7 are reported for the entire corpus, not just the training set.

high-complexity candidates (g–i), occur only a handful of times (h–i are hapaxes) and are assigned vanishing probabilities by the model (cf. Figure 8).

	Example line	HEAVY⇒S 0.86	STRESS⇒(HEAVY⇒S) 1.12	STRESS⇒(S⇒HEAVY)$_{cad}$ 2.62	HEAVY⇒S$_{cad}$ 0.64	STRESS⇒(HIGH⇒S) 1.40	*H*	*frequency*
a.	tô je Mârko ‖ pòslušao mâjku (II, 72:14)						0	1,042
b.	nȉje vȅćē ‖ čȕdo nàstanulo (II, 39:2)	*					−0.86	695
c.	kònopcem mu ‖ savézao rûke (IV, 13:31)	*	*				−1.98	166
d.	nè bi l’ cȁra ‖ u živòtu nàšli (IV, 13:31)			*			−2.62	81
e.	ùdri, brȁte, ‖ ako bȍljē mȍžeš (II, 30:260)	*		*			−3.48	39
f.	nȕt’, dȉvere, ‖ zlâtnī môj pȑstene (I, 342:89)	*				**	−3.66	16
g.	tȅr prìsīcā ‖ dvâ zlâtnā gájtana (I, 342:65)	***	**			*	−6.22	2
h.	ȅto Ćírē, ‖ a nêmā svȁtōvā (III, 76:130)	***	*		*	**	−7.14	1
i.	da jâ vȉdīm ‖ štȁ ȉmā u gòri (III, 47:209)	**	*	*		**	−8.26	1

Table 7: Correlation between Harmony and line type frequency.

6.1.3 Generalization

We now assess how well the grammar established in §6.1.2 generalizes to new, unseen lines using holdout validation. Specifically, we use the constraint weights learned in training (Table 6) to predict line type probabilities for the test set. As shown in Figure 9, the agreement between the observed and predicted probabilities within the test set is high ($r^2 = 0.941$), nearly matching the training fit ($r^2 = 0.952$; Figure 8). This indicates that the model does not overfit the training data: The grammar captures the regularities of the corpus and generalizes comparably well to unseen lines.

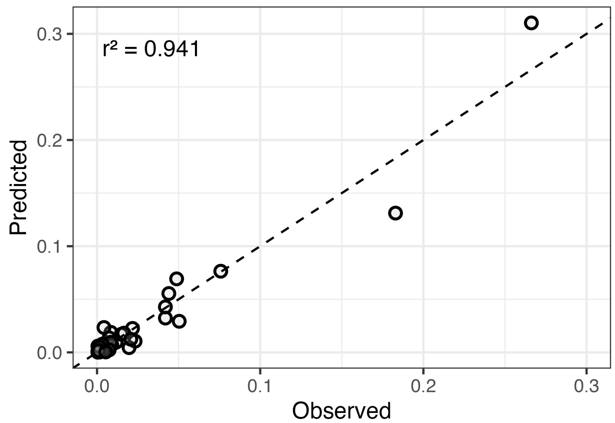


Figure 9: Model fit to test data.

As a second test, we evaluate the model on line types argued by previous scholarship to be unmetrical. A metrical grammar must not only generate the attested line types, but also rule out unmetrical structures (Hayes et al., 2012). Our model was trained exclusively on attested lines (one half of the dataset used for the statistical analysis), so this method tests the model's ability to rule out structures that should be impossible. Consider the lines in (23), which instantiate structures that violate Jakobson (1966)'s metrical constants. All lines in (23) contravene the zeugma, whereby "the fourth and the tenth syllable belong to the same word unit as the third and ninth syllable, respectively" (Jakobson, 1966, 418). Second, in the epic decasyllable, at least one word boundary must occur before an odd position. All disyllables in (23) directly violate this principle, thereby misaligning multiple pitch accents. As shown, our model assigns extremely high penalties to these lines, correctly driving their predicted probabilities towards zero.[13]

(23) a. pa krênū tȁd ‖ da brȁta nâdjē ôn
so moved then that brother find he
'So he moved then to find (his) brother.' ($H = -13.58$)

b. kad tô čû ôn ‖ tȁd sîdjē ȍpēt tû
when that heard he then descends again there
'When he heard that, he descended there again.' ($H = -15.56$)

c. ak' nè dōdjēš ‖ da čûvā lêpī Bôg
if not come that saves handsome God
'If you don't come, may dear God save us!' ($H = -11.64$)

[13]Lines (23a–23b) are artificially created to violate Jakobson's constants, while (23c) is a permutation of a real line produced by the legendary guslar Avdo Medjedović (*The Wedding of Vlahinjić Alija*, sung version, 2,486; Parry, 1980). The attested counterpart of (23c) (*ak' ne dodješ, da Bog lepi čuva*) is less marked but sill metrically complex ($H = -6.60$) due to a long sequence of heavy syllables; the poet chose the order that maps these heavies in the least costly fashion.

6.2 Metrical constraints and prose baseline

The foregoing discussion invites the question of whether the observed line type frequencies fall out from true metrical preferences or reflect ordinary-language phonology. We use two methods to assess constraint contribution in meter relative to an ordinary-language baseline.

First, following Hayes and Schuh (2019), we analyze line type frequencies in our prose sample as a function of their performance on metrical constraints. The results point to systematic differences between the epic and prose.

Violation-free lines account for merely 12% of the prose sample, compared to 28% in the epic. Violations of individual constraints also show remarkable contrasts. STRESS⇒(HEAVY ⇒STRONG) is violated at least once in nearly one-third of prose lines (32%), but in only 14% of epic lines (Fisher's exact test; OR = 2.89, $p < 0.0001$). Multiple violations are also more common in prose: 5% of prose lines violate this constraint twice or more, compared to just 0.04% in the epic. Similarly, STRESS⇒(HIGH⇒STRONG) is violated at least once in 44% of prose lines, but only one-quarter of epic lines (OR = 2.37, $p < 0.0001$). Multiple violations are found in 12% of prose lines versus 3% in the epic. Perhaps more striking is the difference in STRESS⇒(STRONG⇒HEAVY)$_{\text{cadence}}$, which is violated by 29% of prose lines but only 6% of epic lines (OR = 6.13, $p < 0.0001$).

In addition, we implement Henriksson (2022)'s method, which incorporates prose baselines into Maxent models to tease apart true metrical effects from ordinary-language phonology. We fit the model to our prose corpus and obtain constraint weights as an estimate of their baseline contribution. We then refit the model to the full epic corpus K times (K = number of constraints; 5 in our case). On each iteration i, the default Gaussian prior (which we set to $\mu = 0, \sigma = 100$) is replaced for the ith constraint by its baseline weight obtained from the prose model (μ_i = baseline weight, $\sigma_i = 0.01$), while all other constraints retain the default prior.[14] The low σ value strongly penalizes deviations from μ, nearly reducing the constraint's contribution to its baseline effect. The procedure suppresses one constraint at a time rather than all at once to isolate the effects of each individual constraint.

We then compare each of the K models with constraint-specific priors to the full epic model (with uniform default priors for all constraints). We use the BIC difference between models with constraint-specific priors and the full model to assess whether individual constraints contribute to the metrical grammar beyond their baseline effects. The results are shown in Table 8.[15]

[14] Henriksson (2022) sets $\sigma^2 = 0.001$ (i.e. $\sigma = 0.03$) for the suppressed constraints (because `maxent.ot` uses σ rather than σ^2 (Mayer et al., 2024, 25), we report the former). Our choice of $\sigma = 0.01$ imposes a somewhat stronger penalty for deviations from the baseline weight. Henriksson's default priors are $\mu = 0, \sigma = 1$ (vs. our 100). In our models, we found the difference between $\sigma = 100$ and $\sigma = 1$ to be negligible with $\mu = 0$.

[15] The slight differences in constraint weights between the model in Table 6 and the epic model in Table 8 reflect the fact that the former was fit to one half of the metrical corpus, whereas the latter was fit to the full corpus.

Constraint	weight (epic)	weight (prose)	weight difference	ΔBIC
Heavy⇒S	0.86	0.76	0.1	15.79
Stress⇒(Heavy⇒S)	1.07	0.18	0.89	390.99
Stress⇒(S⇒Heavy)$_{\text{cadence}}$	2.59	0.77	1.82	1,131.4
Heavy⇒S$_{\text{cadence}}$	0.57	0.14	0.43	81.1
Stress⇒(High⇒S)	1.38	0.84	0.54	288.76

Table 8: Constraint contribution in the epic vs. prose model: weight difference and BIC comparison.

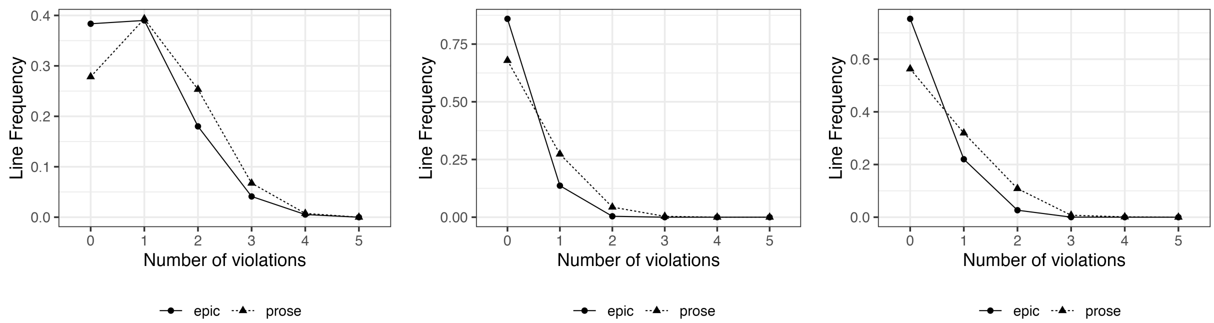
As Table 8 shows, all five constraints receive more weight in the epic than in prose model. The weight differences are modest for some constraints (e.g. Heavy⇒Strong), but substantial for others: Stress⇒(Strong⇒Heavy)$_{\text{cadence}}$ more than triples its baseline weight and Stress⇒(Heavy⇒Strong) increases nearly sixfold. This confirms that the epic regulates weight in stressed syllables more stringently than expected given the ordinary-language patterns (§4.2.1). Stress⇒(High⇒S) is individually the strongest constraint in prose, which serves as a reminder that stressed High-toned syllables are inherently biased towards odd positions. Even so, much of the tonal effect is exclusive to the meter. The BIC comparisons show that suppressing any constraint to its baseline contribution leads to a significant loss in terms of model fit, confirming that each constraint has an independent metrical contribution.

In summary, compared to prose, the epic consistently overrepresents metrically desirable line types and underrepresents marked ones. A Henriksson (2022)-style analysis shows that each constraint in our final model plays an active role in the meter, not reducible to its baseline, ordinary-language effect.

6.3 Cumulativity and the OT-HG debate

Our metrical data are rife with cumulative interactions, whereby multiple metrical factors, or multiple instances of a single factor, jointly contribute to line type probability (cf. Jäger and Rosenbach, 2006; Pater, 2009; Albright, 2012; Breiss, 2020). We observed both counting and ganging-up cumulativity. In counting cumulativity, line types get incrementally more marked with multiple violations of a single constraint (Kawahara, 2020; Kawahara and Breiss, 2021; Kim, 2022). In ganging, coincident violations of multiple independent constraints give rise to a more severe penalty than single constraint violations (Pater, 2016; Smith and Pater, 2020; Breiss and Albright, 2022).

Counting cumulativity manifests in the fact that line type frequencies decline with multiple violations of the same constraint. That multiple violations are rarer than single or no violations is hardly surprising. For example, for Stress⇒(Heavy⇒S) to be violated multiple times by the same line, the line needs to contain multiple instances of stressed heavies, which is already not very common, and have them all occupy weak positions. Indeed, line frequencies drop with multiple violations in both epic and prose (Figure 10).



(a) Heavy⇒S (b) Stress⇒(Heavy⇒S) (c) Stress⇒(High⇒S)

Figure 10: Counting cumulativity: line type frequency (y-axis) as a function of the number of constraint violations (x-axis) in epic vs. prose (point shape and line type).

However, as we alluded to in §6.2, multiple violations are rarer in the meter than in prose. All constraints that allow multiple violations per line exhibit a more concave curve in the epic.[16] The ratio of epic to prose decreases monotonically with violations (Figure 11), which indicates that violating constraints multiple times has a more detrimental effect on probability in the epic than in prose.

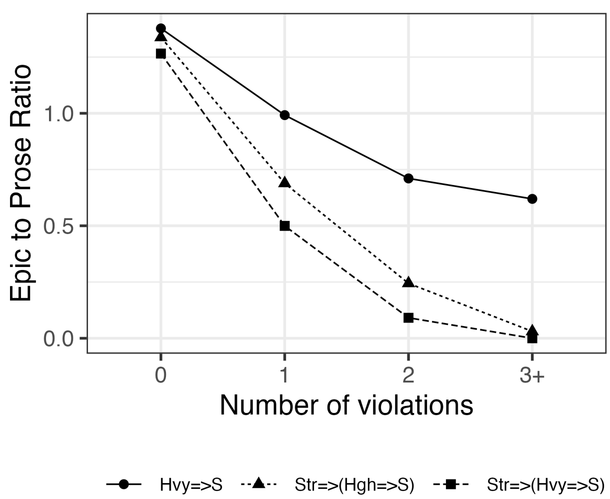


Figure 11: Epic to prose ratio (vertical axis) by the number of violations (horizontal axis) for three constraints (line type and point shape).

We also find pervasive ganging-up cumulativity. Coincident violations of multiple independent constraints are generally rarer than individual violations, as Figure 12 demonstrates.[17]

[16] Heavy⇒Strong$_{\text{cadence}}$ and Stress⇒(Strong⇒Heavy)$_{\text{cadence}}$ can be violated only once per line.

[17] We analyze constraint pairs where independent violations of both conjuncts were possible, therefore excluding pairs in a superset relation.

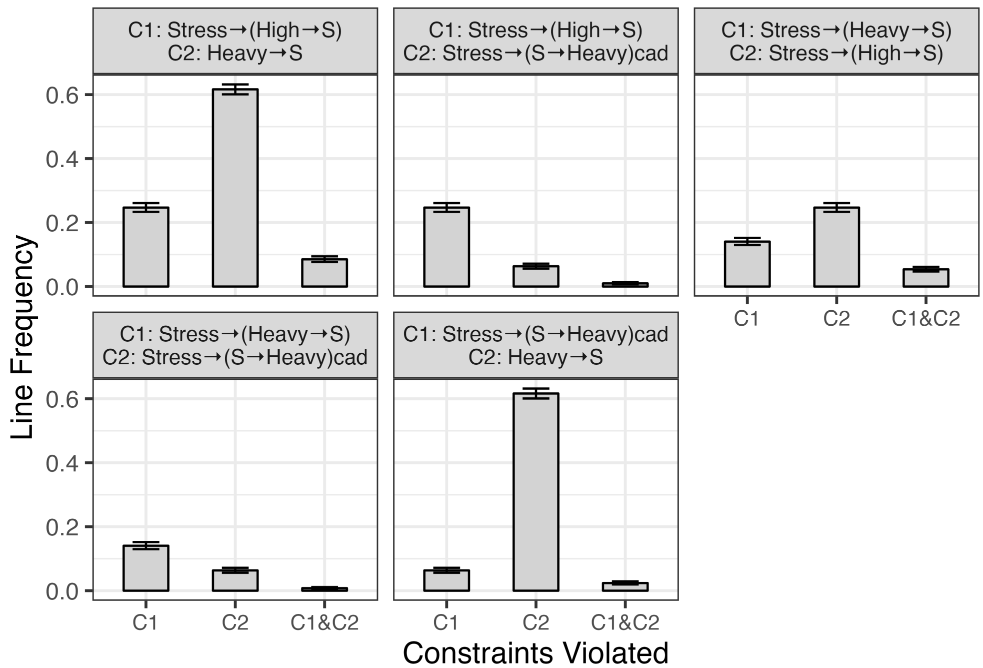


Figure 12: Each facet shows a pair of constraints (C1 and C2). The three bars represent lines that violate just C1, just C2, and both C1 and C2, respectively. Bar height shows the frequency of each violation profile.

We subsequently tested whether the joint effects of ganging constraints match the combinations of their independent contributions (linearity) or not (nonlinearity). In nonlinear interactions, two or more constraints add up to less (sublinearity) or more (superlinearity) than the expected combined effect of their independent contributions.

We estimate expected frequencies of doubly-violating lines from individual violation rates under the assumption of statistical independence (cf. Hayes et al., 2012; Breiss and Albright, 2022) and compare these frequencies with their observed frequencies to obtain O/E ratios. An O/E ratio close to 1 indicates linearity, O/E $>$ 1 indicates sublinearity (joint violations are less detrimental than expected), while O/E $<$ 1 indicates superlinearity (joint violations are more detrimental than expected).

To illustrate, since Stress⇒(Heavy⇒Strong) and Stress⇒(High⇒Strong) are violated in 14% and 25% of lines, respectively, we predict $0.14 \times 0.25 = 3.5\%$ joint violations, but the observed rate is 5.4%, suggesting weak sublinearity (see Table 9).

constraint pair	expected	observed	O/E
STRESS⇒(HEAVY⇒S)& STRESS⇒(HIGH⇒S)	3.5%	5.4%	1.55
STRESS⇒(HEAVY⇒S)& STRESS⇒(S⇒HEAVY)$_{\text{cadence}}$	0.09%	0.08%	0.89
STRESS⇒(HIGH⇒S)& STRESS⇒(S⇒HEAVY)$_{\text{cadence}}$	1.6%	1%	0.63
STRESS⇒(S⇒HEAVY)$_{\text{cadence}}$& HEAVY⇒S	3.9%	2.4%	0.56
STRESS⇒(HIGH⇒S)& HEAVY⇒S	15.2%	8.5%	0.61

Table 9: Expected and observed rates of doubly-violating lines.

In Maxent, local constraint conjunction (Smolensky, 2006) is sometimes invoked to capture extreme cases of superlinearity (Shih, 2017). For all five constraint pairs in Table 9, we tested whether adding conjoined constraints improved model fit. None of the conjoined constraints passed the likelihood ratio test (Table 10), suggesting that our current model does not need interaction terms to capture observed cumulative interactions.

conjunction	$\chi^2(1)$	p
STRESS⇒(HEAVY⇒S) & STRESS⇒(HIGH⇒S)	0.00002	0.996
STRESS⇒(HEAVY⇒S) & STRESS⇒(S⇒HEAVY)$_{\text{cadence}}$	0.00001	0.997
STRESS⇒(HIGH⇒S) & STRESS⇒(S⇒HEAVY)$_{\text{cadence}}$	1.02	0.31
STRESS⇒(S⇒HEAVY)$_{\text{cadence}}$ & HEAVY⇒S	0.000005	0.998
STRESS⇒(HIGH⇒S) & HEAVY⇒S	0.00002	0.997

Table 10: Likelihood ratio tests for conjoined constraints.

The identification of both ganging-up and counting cumulativity in BCMS meter has implications for phonological theory (cf. Breiss, 2020; Kawahara and Breiss, 2021 for other cases where the two types of cumulativity coexist). In HG, candidates' well-formedness is evaluated in a holistic fashion: All constraint violations, as long as the constraint violated has non-zero weight, contribute to the candidate's Harmony. HG thus predicts that constraint violations accumulate both within individual constraints (counting cumulativity) and across constraints (ganging-up cumulativity) (Jäger and Rosenbach, 2006; Pater, 2009).

By contrast, OT employs "fast and frugal" assessment (Kawahara and Breiss, 2021): The winner is selected based on candidates' performance on the top-ranking constraint(s). Lower-ranked constraints cannot affect the outcome, even when they overwhelmingly favor a candidate that fares poorly on the highest-ranking constraint, as in (24). Candidate 2 beats Candidate 1 in (24) because constraint 𝔸 individually outranks every other constraint.

(24) No cumulativity effects in OT

	𝔸	𝔹	ℂ	𝔻	𝔼
a. Candidate 1	*!				
b. ☞ Candidate 2		***	**	*	*

Example (24) shows that base OT rules out both ganging-up and counting cumulativity. While OT can accommodate ganging-up cumulativity by adopting local conjunctions (e.g., in (24), Candidate 1 can prevail against Candidate 2 if the conjoined constraint 𝔻&𝔼 ranks over 𝔸), its predictions regarding counting cumulativity effects are particularly problematic. In OT, violation counts matter only in restricted scenarios: They are relevant only when competing candidates incur different numbers of violations of the highest-ranking constraint that distinguishes between them (25). In (25), both candidates violate the top-ranking constraint 𝔸 at least twice; Candidate 1 incurs the third, fatal, violation.

(25)

	𝔸	𝔹	ℂ
a. Candidate 1	***!		*
b. ☞ Candidate 2	**		*

Crucially, violation counts in OT matter only for the highest-ranking active constraint (25), but are completely irrelevant for lower-ranked constraints (24). This prediction sharply distinguishes OT from HG, where each additional violation of a constraint makes the candidate less harmonic, irrespective of how strong the constraint is individually (26).

(26)

	𝔸 2	𝔹 1	ℂ 0.4	H
a. ☞ Candidate 1	*	*		−3
b. Candidate 2	*	*	**	−3.8
c. Candidate 3	*		****	−3.6

The BCMS meter strongly supports HG's predictions. Constraint violations not only interact cumulatively, but cumulativity effects are pervasive: every constraint that permits multiple violations per line exhibits counting cumulativity (Figure 11), and ganging occurs for every pair of constraints that can be violated independently (12). In other words, multiple violations of the same constraint and coincident violations of distinct constraints result in more severe penalties for lines. These findings favor HG's holistic assessment of candidate well-formedness over OT's "fast and frugal" selection (Kawahara and Breiss, 2021), adding to the growing body of evidence for HG as the model of phonology (Zuraw and Hayes, 2017; Breiss, 2020; Smith and Pater, 2020; Hayes, 2022, 2023).

7 Conclusion

This article identifies a complex pattern of interactive prominence mapping in poetic meter. The BCMS epic decasyllable is a hybrid meter in which both weight and tone regulation are modulated by stress. Specifically, stressed heavies and stressed High-toned syllables are avoided in weak positions throughout the line. Moreover, in the strong position of the cadence, stressed syllables are preferably heavy. Unstressed syllables, by contrast, are largely unregulated, even if they are heavy and/or High-toned.

The parallel interactive mappings observed in the BCMS folk meter expand the empirical coverage of Ryan (2017)'s typology of hybrid meters. Whereas other cases of interactive mapping involve stress and weight, the epic decasyllable displays both stress-weight and stress-tone interaction, the latter being previously undocumented. This is also the first documented case of tone-sensitivity in Indo-European metrical traditions, albeit mediated by stress. Although Sanskrit and Ancient Greek have pitch accents, the meter ignores them.

What sets the present account of the epic decasyllable apart from its predecessors (Jakobson, 1966; Batinić, 1975; Zec, 2008) is the use of inferential statistics as well as controls for confounding factors and ordinary-language distributional skews. We confirm that regulation effects in the meter reflect true metrical preferences, not reducible to extraneous factors.

We provide a Maxent analysis of the epic decasyllable, which uses rigorous principles for the determination of metricality, probability estimation, and assessing competing grammars (Hayes et al., 2012; Hayes, 2023). Compared to prose, the meter is shown to overrepresent metrically well-formed line types and to avoid marked structures. We also identify ubiquitous cumulative patterns, that is, additive effects of constraint violations both within and across constraints, in line with the predictions of HG.

HG (specifically, Maxent) is also useful in this case because it allows the incorporation of priors (i.e. default constraint weights, as specified by μ in the regularization term). Because meter is superimposed on natural language, such priors can be set to their natural-language levels in order to isolate and model the specific contribution of meter. Priors are thus expedient not just for implementing learning biases (Wilson, 2006; Hayes and White, 2015), but also for testing or implementing an additional layer of strictness (here, meter) within the same language (see also Henriksson 2022).

Competing interests

The authors have no competing interests to declare.

References

Albright, Adam. 2012. Additive markedness interactions in phonology. Unpublished manuscript, MIT.

Allen, W Sidney. 1973. *Accent and rhythm*. Cambridge: Cambridge University Press.

Arlot, Sylvain, and Alain Celisse. 2010. A survey of cross-validation procedures for model selection. *Statistics Surveys* 4.40–79.

Arnold, Edward Vernon. 1905. *Vedic metre in its historical development*. Cambridge: University Press.

deCastro Arrazola, Varuṇ. 2018. Testing the robustness of final strictness in verse lines. *Studia Metrica et poetica* 5.55–76.

Bailey, James. 1975. *Toward a statistical analysis of english verse. the iambic tetrameter of ten poets*. Lisse: Peter de Ridder.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67.1–48.

Batinić, Pavle. 1975. *The meter of the Serbo-Croatian oral epic decasyllable*. Doctoral Dissertation, Columbia University.

Bickmore, Lee S. 1995. Tone and stress in Lamba. *Phonology* 12.307–341.

Blumenfeld, Lev. 2016. Generative metrics: An overview. *Language and Linguistics Compass* 10.413–430.

Breiss, Canaan. 2020. Constraint cumulativity in phonotactics: Evidence from artificial grammar learning studies. *Phonology* 37.551–576.

Breiss, Canaan, and Adam Albright. 2022. Cumulative markedness effects and (non-) linearity in phonotactics. *Glossa* 7.

de Lacy, Paul. 2002. The interaction of tone and stress in Optimality Theory. *Phonology* 19.1–32.

Della Pietra, Stephen, Vincent Della Pietra, and John Lafferty. 1997. Inducing features of random fields. *IEEE transactions on pattern analysis and machine intelligence* 19.380–393.

Devroye, Luc, and Terry Wagner. 1979. Distribution-free performance bounds for potential function rules. *IEEE Transactions on Information Theory* 25.601–604.

Fabb, Nigel, and Morris Halle. 2008. *Meter in poetry: A new theory*. Cambridge: Cambridge University Press.

Flemming, Edward. 2021. Comparing MaxEnt and Noisy Harmonic Grammar. *Glossa: a journal of generative linguistics* 6.

Foley, John Miles. 1993. *Traditional oral epic: The Odyssey, Beowulf, and the Serbo-Croatian return song*. Berkeley: University of California Press.

Gasparov, Mikhail L. 1980. Quantitative methods in Russian metrics: Achievements and prospects. *Russian poetics in translation* 7.1–19.

Gasparov, Mikhail L. 1987. Probability model of verse (English, Latin, French, Italian, Spanish, Portuguese). *Style* 21.

Goldsmith, John. 1987. Tone and accent, and getting the two together. *Annual Meeting of the Berkeley Linguistics Society*, volume 13, 88–104.

Goldwater, Sharon, and Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. *Proceedings of the workshop on variation within Optimality Theory*, 111–120.

Golston, Chris, and Tomas Riad. 2000. The phonology of Classical Greek meter. *Linguistics* 38.

Gordon, Matthew. 2006. *Syllable weight: phonetics, phonology, typology*. New York: Routledge.

Gordon, Matthew. 2023. The phonetic basis for tone-stress interactions: A cross-linguistic study. *Representing phonological detail, part II: Syllable, stress, and sign*, ed. by Jeroen van de Weijer, 171–190. Berlin: Walter de Gruyter.

Gunkel, Dieter C, and Kevin Ryan. 2011. Hiatus avoidance and metrification in the Rigveda. *Proceedings of the 22nd Annual UCLA Indo-European Conference*, ed. by Stephanie W. Jamison, H. Craig Melchert, and Brent Vine, 53–68. Bremen: Hempen.

Hall, Daniel Currie. 2006. Modeling the linguistics-poetics interface. *Formal approaches to poetry: Recent developments in metrics*, ed. by Nila Dresher, B. Elan Friedberg, 233–49. Berlin: Mouton de Gruyter.

Halle, Morris, and Samuel Jay Keyser. 1966. Chaucer and the study of prosody. *College English* 28.187–219.

Halle, Morris, and Samuel Jay Keyser. 1971. Theory of prosody, continued: Illustration and defense of a theory of the iambic pentameter. *College English* 33.154–176.

Hanson, Kristin, and Paul Kiparsky. 1996. A parametric theory of poetic meter. *Language* 72.287–335.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. *The elements of statistical learning*. New York: Springer.

Hayes, Bruce. 1983. A grid-based theory of English meter. *Linguistic Inquiry* 14.357–393.

Hayes, Bruce. 1988. Metrics and phonological theory. *Linguistics: The cambridge survey: Volume 2, linguistic theory: Extensions and implications*, ed. by Frederick J. Newmeyer, 220–249. Cambridge: Cambridge University Press.

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago and London: University of Chicago Press.

Hayes, Bruce. 2022. Deriving the wug-shaped curve: A criterion for assessing formal theories of linguistic variation. *Annual Review of Linguistics* 8.473–494.

Hayes, Bruce. 2023. A brief introduction to meter. Talk at M@100, MIT.

Hayes, Bruce, and Donka Minkova. 2023. A MaxEnt analysis of the meter of Beowulf. Poster at M@100, MIT.

Hayes, Bruce, and Claire Moore-Cantwell. 2011. Gerard Manley Hopkins' sprung rhythm: Corpus study and stochastic grammar. *Phonology* 28.235–282.

Hayes, Bruce, and Russell G Schuh. 2019. Metrical structure and sung rhythm of the Hausa rajaz. *Language* 95.e253–e299.

Hayes, Bruce, and James White. 2015. Saltation and the p-map. *Phonology* 32.267–302.

Hayes, Bruce, and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic inquiry* 39.379–440.

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

Henriksson, Erik. 2022. *Greek meter: an approach using metrical grids and maxent*. Doctoral Dissertation, Helsingin yliopisto.

Inkelas, Sharon, and Draga Zec. 1988. Serbo-Croatian pitch accent: the interaction of tone, stress, and intonation. *Language* 64.227–248.

Ivić, Pavle. 1958. *Die serbokroatischen Dialekte. ihre Struktur und Entwicklung*. The Hague: Mouton.

Jäger, Gerhard, and Anette Rosenbach. 2006. The winner takes it all—almost: Cumulativity in grammatical variation. *Linguistics* 44.937–971.

Jakobson, Roman. 1966. Slavic epic verse: studies in comparative metrics. *Selected writings*, volume 4, 414–463. The Hague & Paris: Mouton.

Janson, Tore. 1975. *Prose rhythm in medieval latin from the 9th to the 13th century*. Stockholm: Almqvist & Wiksell International.

Kager, René. 1989. *A metrical theory of stress and destressing in English and Dutch*. Doctoral Dissertation, Utrecht University.

Karadžić, Vuk Stefanović. 1824. Predgovor. *Narodne srpske pjesme, skupio ih i na svijet izdao vuk stef. karadžić. knjiga prva, u kojoj su različne ženske pjesme*. Leipzig: Breitkopf & Härtel.

Karadžić, Vuk Stefanović. 1841–1862. *Srpske narodne pjesme. Knjiga prva, u kojoj su razlučne ženske pjesme (1841). Knjiga druga, u kojoj su pjesme junačke najstarije (1845). Knjiga treća, u kojoj su u kojoj su pjesme junačke srednjijeh vremena (1846). Knjiga četvrta, u kojoj su pjesme junačke novijih vremenea o vojevanju za slobodu (1862).* Vienna: U štampariji jermenskoga manastira.

Karadžić, Vuk Stefanović. 1852. *Srpski rječnik istumačen njemačkijem i latinskijem riječima*. Vienna.

Kawahara, Shigeto. 2020. A wug-shaped curve in sound symbolism: The case of Japanese Pokémon names. *Phonology* 37.383–418.

Kawahara, Shigeto, and Canaan Breiss. 2021. Exploring the nature of cumulativity in sound symbolism: Experimental studies of pokémonastics with english speakers. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 12.

Kim, Seoyoung. 2022. A maxent learner for super-additive counting cumulativity. *Glossa* 7.

Kiparsky, Paul. 1968. Metrics and morphonemics in the Kalevala. *Studies presented to Roman Jakobson by his students*, ed. by Charles E Gribble, 137–148. Cambridge, MA: Slavica Publishers.

Kiparsky, Paul. 1975. Stress, syntax, and meter. *Language* 51.576–616.

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.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Harmonic grammar—a formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. *Proceedings of the twelfth annual conference of the Cognitive Science Society*, 884–891.

Lehiste, Ilse, and Pavle Ivić. 1986. *Word and sentence prosody in serbocroatian*. Cambridge, MA: MIT Press.

Lord, Albert B. 1960. *The singer of tales*. Cambridge, MA: Harvard University Press.

Lubotsky, Alexander. 1995. Accentuation in the technique of the Vedic poets. *Studies in poetics: commemorative volume krystyna pomorska*, ed. by E. Semeka-Pankratov, 515–534. Columbus, Ohio: Slavica Publishers.

Maas, Paul. 1962. *Greek metre*. Oxford: Clarendon Press.

Macdonell, Arthur Anthony. 1910. *Vedic grammar*. Strassburg: Verlag von Karl J. Trübner.

Maretić, Tomo. 1901. O gradji našega narodnog deseterca. *Nastavni vjesnik* 10.145–175.

Maretić, Tomo. 1907. Metrika narodnih naših pjesama. *Rad Jugoslavenske akademije znanosti i umjetnosti* 168.1–112.

Matić, Svetozar. 1964. *Naš narodni ep i naš stih.*. Novi Sad: Matica srpska.

Mayer, Connor, Adeline Tan, and Kie Ross Zuraw. 2024. Introducing `maxent.ot`: an R package for Maximum Entropy constraint grammars. *Phonological Data and Analysis* 6.1–44.

McPherson, Laura, and Kevin M Ryan. 2018. Tone-tune association in Tommo So (Dogon) folk songs. *Language* 94.119–156.

Nikolić, Berislav. 1970. *Osnovi mladje novoštokavske akcentuacije*. Beograd: Institut za srpskohrvatski jezik.

Parry, Milman. 1980. *Serbo-Croatian heroic songs. Volume VI: Ženidba Vlahinjić Alije, Osmanbeg Delibegović i Pavičević Luka. Kazivao i pjevao Avdo Medjedović*. Cambridge, MA: Harvard University Press. Collected by Milman Parry. Edited with prolegomena and notes by David E. Bynum.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive science* 33.999–1035.

Pater, Joe. 2016. Universal grammar with weighted constraints. *Harmonic grammar and harmonic serialism*, ed. by Joe Pater and John McCarthy, 1–46. London: Equinox.

Prince, Alan. 1990. Quantitative consequences of rhythmic organization. *Papers from the 26th Regional Meeting of the Chicago Linguistic Society: Volume 2: The Parasession*, ed. by Michael Ziolkowski, Manuela Noske, and Karen Deaton, 355–398.

Prince, Alan, and Paul Smolensky. 2004. *Optimality theory: Constraint interaction in generative grammar*. John Wiley & Sons.

R Core Team. 2023. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raven, David S. 1962. *Greek metre: an introduction*. London: Faber & Faber.

Ružić, Žarko. 1975. *Srpski jamb i narodna metrika*. Beograd: Institut za književnost i umetnost.

Ryan, Kevin M. 2011. Gradient syllable weight and weight universals in quantitative metrics. *Phonology* 28.413–454.

Ryan, Kevin M. 2016. Phonological weight. *Language and Linguistics Compass* 10.720–733.

Ryan, Kevin M. 2017. The stress–weight interface in metre. *Phonology* 34.581–613.

Ryan, Kevin M. 2019. *Prosodic weight: Categories and continua*. Oxford: Oxford University Press.

Schwarz, Gideon. 1978. Estimating the dimension of a model. *The annals of statistics* 6.461–464.

Selkirk, Elisabeth. 1996. The prosodic structure of function words. *Sugnal to syntax: Bootstrapping from speech to grammar in early acquisition*, ed. by James L. Morgan and Katherine Demuth. Mahwah, NJ: Erlbaum.

Shih, Stephanie S. 2017. Constraint conjunction in weighted probabilistic grammar. *Phonology* 34.243–268.

Smith, Brian W, and Joe Pater. 2020. French schwa and gradient cumulativity. *Glossa: a journal of general linguistics* 5.

Smith, Liam, and Arto Tapani Anttila. 2025. Metrical tension and prose cadence. *Proceedings of the Annual Meetings on Phonology*, volume 1. University of Massachusetts Amherst Libraries.

Smolensky, Paul. 2006. Optimality in phonology II: Harmonic completeness, local constraint conjunction, and feature domain markedness. *The harmonic mind: From neural computation to optimality-theoretic grammar. vol. 2: Linguistic and pholosophical implications.*, ed. by Paul Smolensky and Géraldine Legendre, 27–160. Cambridge, MA: MIT Press.

Tarlinskaja, Marina. 1976. *English verse: theory and history*. The Hague and Paris: Mouton.

Tarlinskaja, Marina G, and Lilya M Teterina. 1974. Verse–prose–metre. *Linguistics* 12.63–86.

Vaillant, André. 1932. Les Chants épiques des Slaves du Sud. *Revue des cours et conférences* 4.309–326.

Werle, Adam. 2009. *Word, phrase, and clitic prosody in Bosnian, Serbian, and Croatian*. Doctoral Dissertation, University of Massachusetts Amherst.

West, Martin Litchfield. 1987. *Introduction to Greek metre*. Oxford: Clarendon Press.

Wilson, Colin. 2006. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30.945–982.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Zec, Draga. 2000. O strukturi sloga u srpskom jeziku. *Južnoslovenski filolog* 56.435–448.

Zec, Draga. 2005. Prosodic differences among function words. *Phonology* 22.77–112.

Zec, Draga. 2008. The prosodic word as a unit in poetic meter. *The nature of the word: Essays in honor of Paul Kiparsky*, ed. by K. Hanson and S. Inkelas, 63–94. Cambridge, MA: MIT Press.

Zsiga, Elizabeth, and Draga Zec. 2013. Contextual evidence for the representation of pitch accents in Standard Serbian. *Language and speech* 56.69–104.

Zuraw, Kie, and Bruce Hayes. 2017. Intersecting constraint families: an argument for Harmonic Grammar. *Language* 93.497–548.