Authors' manuscript (August 2026), accepted with minor revisions at *Language*. Final version to follow.

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

Aljoša Milenković[1] and Kevin Ryan[1]

[1]Linguistics, Harvard University, Cambridge, MA, USA

Corresponding author: Aljoša Milenković

Received: 11 Oct 2025; **Revised:** ; **Accepted:**

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

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. Most attested line types are harmonically bounded by their competitors, further supporting MaxEnt over frameworks that uphold strict harmonic bounding.

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), 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
Assign a violation for every strong position not containing a heavy syllable.

b. STRONG⇒STRESS$_{\text{cadence}}$
Assign a violation for every strong position in the cadence not containing 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 constraints in (2). Among unstressed and secondary-stressed syllables, weight is less strictly regulated.

(2) STRESS⇒(STRONG⇔HEAVY)

a. STRESS⇒(STRONG⇒HEAVY)
Assign a violation for every light primary-stressed syllable assigned to a strong position

b. STRESS⇒(HEAVY⇒STRONG)
Assign a violation for every heavy primary-stressed syllable assigned to a weak 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, Ryan 2026).

Aside from its contributions to metrical theory and typology, the article also bears on phonological modeling more generally, in three 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, most line types attested in our metrical corpus are harmonically bounded by at least one competitor. This calls for a framework that does not strictly rule out harmonically bounded candidates, such as MaxEnt (Hayes and Moore-Cantwell 2011, Hayes and Schuh 2019). Third, 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 dialect group on which the accentuation of standard BCMS is based, 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). Earlier autosegmental work (Inkelas and Zec 1988, Zec 1999) assumes that in rising-accented syllables, high tone spreads from the posttonic onto the final mora of the stressed syllable, resulting in two consecutive high-toned moras. High spreading is rejected in subsequent work (Zec and Zsiga 2010, 2022, Zsiga and Zec 2013), as it runs counter to the acoustics of NS pitch accents (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	⟨sûnce⟩	[ˈsúun.tsɛ]	‘sun.NOM.SG’
⟨ȁ⟩	short falling	⟨krȁva⟩	[ˈkrá.ʋa]	‘cow.NOM.SG’
⟨á⟩	long rising	⟨rúka⟩	[ˈruu.ká]	‘arm.NOM.SG’
⟨à⟩	short rising	⟨sèstra⟩	[ˈsɛ.strá]	‘sister.NOM.SG’
⟨ā⟩	long unaccented	⟨īgrā⟩	[ˈí.graa]	‘play.PRS.3SG’
⟨a⟩	short unaccented	⟨krȁva⟩	[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.

We follow Inkelas and Zec (1988), Zec (1999), and related autosegmental work in assuming that NS has lexical tone and exhibits tone-driven stress. Tone observes strict culminativity (in the sense of Hyman 2006), in that prosodic words can bear only one, singly-linked high tone. The locus of stress is fully predictable from the position of the lexical high: NS stresses the syllable immediately preceding the lexical high (3a); otherwise (i.e. if the high is word-initial or if the form lacks a high underlyingly), NS stresses the initial syllable (3b) (Zec and Zsiga 2010).

(3) NS stress pattern

a. Stress the syllable preceding the syllable that contains the word's only high-toned mora:

/prɛ.stɔ.lɔ.naa.slɛ́d.niik/	[prɛ.stɔ.lɔ.ˈnaa.slɛ́d.niik]	'crown.prince.NOM.SG'
/in.sti.túut/	[in.ˈsti.túut]	'institute.NOM.SG'

b. Else stress the initial syllable:

/prá.ʋi.ti/	[ˈprá.ʋi.ti]	'make.INF'
/prɔ.pa.dɛ/	[ˈprɔ́.pa.dɛ]	'fail.AOR.2/3SG'

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 (4). The caesura must coincide with a prosodic word boundary, such that it falls neither word-internally nor between a clitic and its host. Unmetrical lines are exceedingly rare (Foley 1993, 88-94).

(4) $\sigma\sigma\sigma\sigma \parallel \sigma\sigma\sigma\sigma\sigma\sigma$

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).

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 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

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

[2] Strictly speaking, Karadžić (1824) discusses 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.

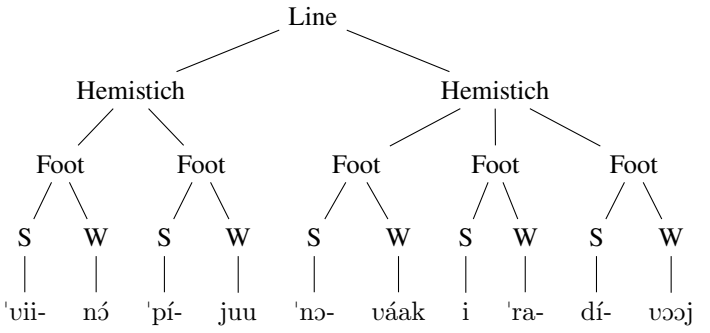


Figure 1.: Epic decasyllable as trochaic pentameter. Translation: 'Novak and Radivoj are drinking wine.' (III, 1:1).

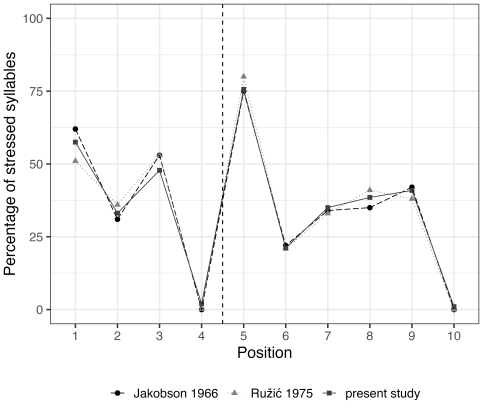


Figure 2.: Proportion of stressed syllables (vertical axis) by position (horizontal axis) and source (line type). The dashed vertical line indicates the caesura.

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 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" quadrisyllabic prosodic words (initially stressed: ˈσσσσ, e.g., [ˈbí.jɛ.li.jɛm] 'white.INS.SG.M/N.DEF') over their "iambic" counterparts (peninitially stressed: σˈσσσ, e.g., [prɛ.ˈpɔ.zná.ti] 'recognize.INF'); 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. Per Ružić (*ibid.*), in prose, peninitially stressed quadrisyllables greatly outnumber (by 3.3 times) initially stressed ones. In meter, by contrast, the two

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

types are approximately equally frequent (Table 4). We replicated Ružić's quadrisyllable test on our corpus and obtained a significant result, but more modest than Ružić's (Fisher's exact test; odds ratio 1.43, $p = 0.01$).

word type	prose	epic decasyllable	prose/epic ratio
ˈσσσσ	3.5%	12.6%	**1:3.6**
σˈσσσ	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 (ˈσσσ) and the other peninitially stressed (σˈσσ), the unmarked order is (5a), locating both stressed syllables in odd positions. The reverse order in (5b) is marked, as it locates both stresses in even positions. In our corpus, the unmarked order (5a) occurs almost twice as often as the marked order (5b) (68 vs. 36 tokens). In our prose sample (see §4.1), by contrast, the same ratio is 15 to 13.

(5) a. Unmarked trisyllable order: σσσσ ‖ ˈσσσ σˈσσ
ˈlɛ́ɛ.pɔ ˈmáaj.ka ‖ ˈɔ́.pra.ʋii bɔg.ˈdaa.ná
'The mother nicely prepared Bogdan.' (I, 723:11)

b. Marked trisyllable order: σσσσ ‖ σˈσσ ˈσσσ
i iz.ˈʋɛ.dí ‖ bi.ˈjɛ.sná ˈʤɔ́.ga.ta
'And bring forth the fierce gray horse.' (IV, 13:104)

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. 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 1862).[4] This corpus is considerably larger than the corpora used in previous studies (cf. 753 lines in Jakobson 1966; 435 in Ružić 1975). Further, our corpus 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 *Serbian Dictionary* (1852), 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, depending on whether proclitics affect the position of stress or are stress-neutral. One option, illustrated in (6), is that all proclitics are part of the stress domain and therefore receive stress in two environments: when their host bears a high on the initial mora (6a), and when the host lacks high-toned moras underlyingly (6b). Alternatively, proclitics may be treated

[4] We used the 1972 reprint edition. The sample included 987 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 4, 11, 13–15).

as external to the stress domain, and therefore never bear stress (7). The dual annotation reflects a split within the NS dialect group, where some dialects follow (6), others follow (7), and a third, hybrid type fluctuates between (6) and (7) (Ivić 1958, Selkirk 1996, Zec 2005, Werle 2009).

(6) Dialects with retraction onto proclitics

a. Host with an initial high tone: stress on the proclitic

/jé.zɛ.ru/	[ˈjé.zɛ.ru]	‘lake.LOC.SG’	[ˈna=jé.zɛ.ru]	‘on=lake.LOC.SG’

b. Toneless host: stress and high tone on the proclitic

/ʊɔ.du/	[ˈʊɔ́.du]	‘water.ACC.SG’	[ˈná=ʊɔ.du]	‘on=water.ACC.SG’

(7) Dialects with no retraction onto proclitics

a. Initial high tone

/jé.zɛ.ru/	[ˈjé.zɛ.ru]	‘lake.LOC.SG’	[na=ˈjé.zɛ.ru]	‘on=lake.LOC.SG’

b. Toneless host

/ʊɔ.du/	[ˈʊɔ́.du]	‘water.ACC.SG’	[na=ˈʊɔ́.du]	‘on=water.ACC.SG’

Following Maretić (1901), Ružić (1975), and Batinić (1975), our analysis adheres to the pattern in (6), treating all proclitics as part of the stress domain. That the meter follows (6) rather than (7) is evidenced 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 onto a proclitic ensures that the otherwise strict ban on stressed monosyllables in weak positions is maintained.

We treat the verbal negation marker *ne* [nɛ] as part of the following prosodic word, despite being orthographically separated from the verb. Thus, there is a discrepancy between the orthographic ‘word-hood’ of *ne* and its prosodic status (8): *ne* is always part of the innermost prosodic word, since it receives stress even in dialects that otherwise treat proclitics as external to the stress domain (Zec 2005); see (7). We accordingly treat (8b) as a trisyllabic prosodic word rather than a monosyllable followed by a disyllable.

(8) a. ⟨glȅdā⟩ [ˈglé.daa] ‘watch.PRS.3SG’

b. ⟨nè gledā⟩ [ˈnɛ=glé.daa] ‘NEG=watch.PRS.3SG’

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).[5] This leaves 30,168 of the original 37,710 syllables.

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 long falling accented syllables (i.e. stressed high-toned heavies) occupy strong positions.

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

[5] The independent factors being that the hemistich-final positions are always filled by word-final syllables, which are never stressed in BCMS (§2.1), and stressed monosyllables are strictly banned in even positions other than 2 and 6.

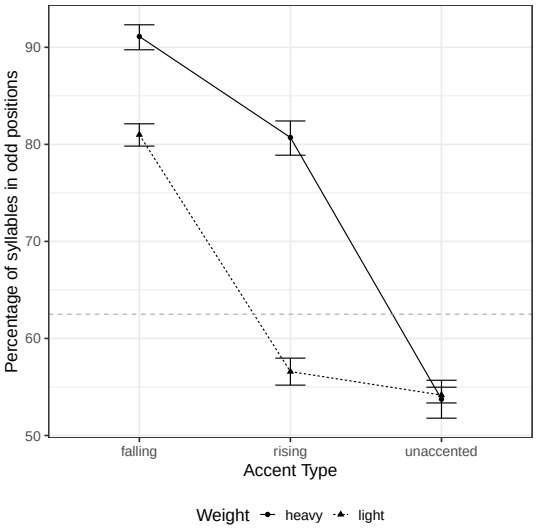


Figure 3.: Proportion of syllables in odd positions (vertical axis) by accent type (horizontal axis) and weight (line type). Error bars represent Wilson 95% confidence intervals. The dashed rule marks the chance level (grand mean).

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.[6] Following Ryan (2026), WORD SHAPE encodes 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.). A few examples of word shape encoding are provided in (9). In each case, X represents the syllable whose properties are being predicted by the model and therefore withheld.

(9)	a.	ˈbɛ.sjɛ́.di.la	X-light-light-light	'speak.PTCP.PST.F.SG'
	b.	ˈɔ́.tuu.rii	light-X-heavy	'push.aside.AOR.3SG'
	c.	ɔd.ˈmaak.níi	light-light-X	'move.away.IMP.2SG'
	d.	ˈna.glɛ́.dax	light-light-X	'see.much.AOR.1SG'

Some control for word shape is necessary because word shape is a confound: The meter regulates where words of various shapes may be localized in the line, and accent and weight are unevenly distributed across word shapes (§4.1). Indeed, if the random effect for shape is removed (all else being equal), the model fit collapses, as indicated by an increase in the Bayesian Information Criterion (BIC; Schwarz 1978): ΔBIC (model with no random effect − full model) = 15,411. At the same time, while the

[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.

details of the encoding just described are somewhat arbitrary, they are also immaterial: Any reasonable operationalization of shape yields the same conclusions.[7]

Random *intercepts* for higher-level units such as poem, book, or singer, by contrast, would be vacuous in this design: The meter fixes five strong positions among each line's eight analyzed positions, so every line—and hence every poem and book—exhibits the same rate of strength ($5 \div 8$). Therefore, when such intercepts are added, their fitted variances are (near-)zero (poem: 0; book: 0.0002). Block-level variation could, however, reside in *slopes*, that is, in the degree to which accent and weight condition strength in one poem or book as opposed to another. The model therefore includes by-poem random slopes for the two accent contrasts and for weight (uncorrelated). Poems ($N = 37$) are individual songs and hence the closest available proxies for individual singers. As Table 5 shows, the fitted slope variances are minuscule. They do not significantly improve fit (likelihood ratio $\chi^2(3) = 4.46$, $p = 0.22$; the same test with by-book rather than by-poem slopes: $\chi^2(3) = 1.95$, $p = 0.58$), and the fixed effects are virtually identical with or without them. In short, prominence mapping is homogeneous across poems and books.

The model results are provided in Table 5. The variance of the word shape random effect is large (SD = 6.95), reflecting the fact that word shapes differ radically (indeed, sometimes categorically) in their localization: 47 of the 68 shapes (6% of syllable tokens) occur exclusively in strong or exclusively in weak positions, and the estimated shape effects are correspondingly bimodal, separating strong-seeking from weak-seeking shapes rather than approximating a normal distribution. The random effect thus serves as a regularized control for shape rather than as a model of a shape population. To confirm that nothing hinges on this idealization (nor on the non-independence of syllables within a line, exactly five of eight of which must be strong), we refit the model as a conditional logistic regression stratified by line, with shape as a fixed factor, which conditions out both issues. The fixed-effect estimates closely track Table 5 (0.18, 0.97, 0.27, 0.73, and −0.68, respectively, all $p < 0.01$).

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 (no stress; tone irrelevant). 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 *post hoc* pairwise comparisons. In unaccented syllables, lights do not differ significantly from heavies ($\beta = -0.06, p = 0.32$). In rising syllables, lights are significantly less strong-aligned than heavies ($\beta = -0.74, p < 0.0001$). Finally, in falling syllables, there is no significant effect of weight ($\beta = 0.03, p = 0.79$).[8]

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 family of sigmoids sometimes described as a "wug shape" (McPherson and Hayes 2016, 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).

[7] For example, one could additionally encode the tonal specifications of non-target syllables. We avoid doing so because target tone is highly correlated with non-target tone within a word. But the results are qualitatively unaffected by this choice: With tone-inclusive shapes (198 levels), all five fixed-effect coefficients of interest remain significant with the same signs (0.37, 1.06, 0.28, 0.55, and −0.82, respectively). The same holds if weight is instead excluded, such that shape encodes only word size and target position (19 levels; coefficients: 0.35, 0.94, 0.28, 0.68, and −0.74, all $p < 0.001$).

[8] 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.

Random effects	*Variance*	*Standard Deviation*		
Word shape (N = 68)				
(Intercept)	48.32	6.95		
Poem (N = 37)				
Rising (vs. Unaccented) slope	0.026	0.16		
Falling (vs. Rising) slope	0.030	0.17		
Heavy slope	0.000	0.00		
Fixed effects	*β*	*Standard Error*	*Wald z*	*p*
(Intercept)	0.75	0.90	0.83	0.41
Accent type (difference coded)				
Rising (vs. Unaccented)	0.29	0.07	4.11	<0.001 ***
Falling (vs. Rising)	1.01	0.08	12.86	<0.001 ***
Weight (baseline: Light)				
Heavy	0.26	0.05	5.01	<0.001 ***
Accent type*Weight interaction				
Rising (vs. Unaccented):Heavy	0.68	0.11	6.07	<0.001 ***
Falling (vs. Rising):Heavy	−0.77	0.14	−5.56	<0.001 ***

Table 5.: Mixed-effects logistic regression model output. Model specification: strong ~ accent_type * heavy + (1 | shape) + (0 + rising_vs_unaccented | poem) + (0 + falling_vs_rising | poem) + (0 + heavy | poem); accent_type is forward-difference coded over unaccented, rising, falling; heavy is 0 for light, 1 for heavy; rising_vs_unaccented = 1 for rising or falling (carrying over the difference coding of the fixed effects); falling_vs_rising = 1 for falling.

The model included random intercepts for word shape, reducing its sensitivity to specific word shapes' idiosyncratic distributions, and by-poem random slopes, which showed prominence mapping to be homogeneous across poems. In this section, we control for potential confounds more directly using baselines of comparison (i.e. observed vs. expected distributions), 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; Foley 1993, 99.) Further, 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.

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. Additionally, 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, which we introduce in the following section.

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, 2026 for scrambled metrical corpora).

4.2.1. Prose sample

As the first control, 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 16 BCMS novels and short story collections by 10 authors, from the late 19th to the early 21st century (titles and authors are listed in Appendix A). An R script retrieved all decasyllabic sentences from the prose corpus. We set sentence boundaries as a period, question mark, or an exclamation mark followed by a capital letter. This definition of sentence units is not without problems, but it was the best possible approximation given the otherwise unprocessed plain-text input. The script assigned length, i.e. syllable count, to each prose sentence. Given that BCMS orthography faithfully represents the pronunciation, and the language has no diphthongs, sentence length was operationalized as the number of vowel characters in the sentence (including syllabic *r*).

From the resulting pool of decasyllabic sentences, we manually extracted sentences with a prosodic word boundary after the fourth syllable, as in (10a), eliminating sentences in which the fourth and fifth syllables belong to the same prosodic word, as in (10b). (Both examples are from Borislav Pekić's novel *Rabies* (*Besnilo*, 1983).) The boundary position matches the epic decasyllable's caesura, one of its metrical constants (Jakobson 1966). This practice follows Gasparov's (1987) suggestion that prose comparanda should comply with the rules of the meter that are independent of those being tested.

(10) Examples of decasyllabic sentences from prose

a. Word boundary after the fourth position (included)
ˈné.kɔ ˈmɔ́ɔ.raa ‖ ˈdɔ́.lɛ da ˈɔ.stá.nɛɛ
'Somebody must remain downstairs.'

b. No word boundary after the fourth position (excluded)
ˈtɔ́ɔ sɛ ˈnɛ=ráa.dii gɔ.ˈʋɔ.rí.li su
'That is not how things are done, they would say.'

Our final prose corpus consists of 801 epic decasyllable-like sentences. Sentences that are structurally parallel to folk verse are difficult to come by in prose. Decasyllabic sentences account for just 2.6% of our prose sample, and of these, only about one-third exhibit the requisite boundary placement (less than 1% of all prose sentences; see Appendix A). The prose sentences were prosodically annotated following the same procedure used for the folk epic corpus, as described in §3.1.

4.2.2. Scrambled metrical corpus

The scrambled metrical corpus was constructed following the "Rigged Veda" method of Gunkel and Ryan (2011), an implementation of the Monte Carlo method (Ryan 2026, 13). 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 using the scrambled corpus). For example, for the real line in (11), the program generates permuted lines such as those in (12).

(11) Real line example
ˈsɔɔ.ʋá ˈsjɛ.díi ‖ na=ˈbúu.kɔ.ʋu ˈpaa.ɲú
'The owl is sitting on the beech stump.' (I, 717:1)

(12) Permuted, nonsensical lines generated by the scrambler

a. ˈsúun.tsɛ ˈlɛ.tíi ‖ ka=ˈŋii.nɔ́.mɛ ˈgráa.du
b. ˈnaa.rɔ́d ˈpá.daa ‖ sa=ˈmlaa.dí.jɛm ˈdáa.nɔm
c. ˈrúu.ku ˈɔ.pɛ́ɛt ‖ na=ˈpɛ́ɛn.ʤɛ.re ˈdaa.jí

4.3. Results

4.3.1. 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.

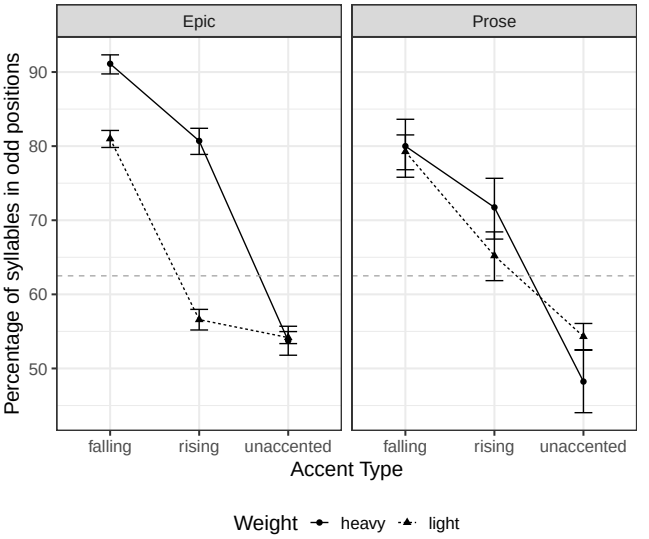


Figure 4.: Proportion of syllables that occur in odd positions (vertical axis), by accent type (horizontal axis) and syllable weight (line type) in the oral epic vs. the epic-like prose sample (facets). The dashed rule indicates chance level.

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.[9] 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 epiphenomenal. 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 ($\beta = 0.18, p = 0.16$), and the contrast is significantly weaker in prose than in the

[9] More precisely, when the regression model from §3.2 is run on the prose sample, with source text ($N = 16$) in place of poem as the block-level grouping factor, weight is weakly significant as a main effect ($\beta = 0.22, p = 0.01$), but more importantly, neither of the interactions between accent type and weight is significant ($\beta = 0.05, p = 0.82$; $\beta = -0.28, p = 0.23$).

epic (difference = 0.83, $p < 0.0001$), while rising vs. unaccented does not differ between the corpora (difference = 0.24, $p = 0.12$).[10] 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 were coded as high), weight, and their interaction, as well as with random intercepts for word shape, and by-poem random slopes for tone and weight.[11] The stressless model substantially underperforms the model with stress, the BIC of the former being higher by 297.

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 though less conclusive at this point, judging by the visualizations, but we present more rigorous evidence supporting it in §4.3.3.

4.3.2. 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'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).

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.

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 (now with 26,397 syllables and 62 word shape levels). In this new model, the main effect of weight disappears: $\beta = -0.025, p = 0.68$. However, the interactions between weight and accent type persist, with the same directionalities: rising (vs. unaccented) * heavy ($\beta = 0.58, p < 0.0001$) and falling (vs. rising) * heavy ($\beta = -0.77, p < 0.0001$), confirming that stress-weight interactions are not confined to the cadence.

[10] Because the epic and prose samples are disjoint, the two models' estimates are independent, so each contrast can be compared across corpora directly ($z = 5.67$ and $z = 1.57$, respectively). A pooled model with corpus interactions yields the same conclusions.

[11] Model specification: strong ~ high * heavy + (1 | shape) + (0 + high | poem) + (0 + heavy | poem); high is coded as 1 for falling and post-rising, 0 otherwise.

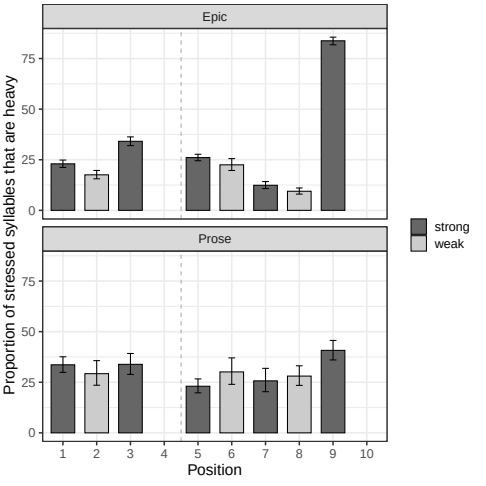


Figure 5.: The proportion of stressed syllables that are heavy (vertical axis) by position (horizontal axis) in epic vs. prose (facets).

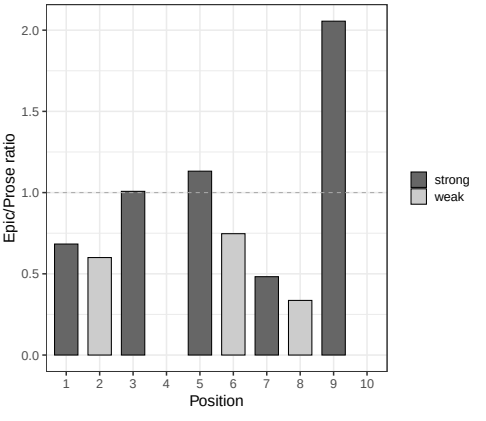


Figure 6.: Epic/prose ratio of stressed heavy syllables (vertical axis) by position (horizontal axis). The dashed horizontal rule is the baseline.

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. 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 near 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 noncanonical noun-adjective order is more likely to be employed (Milenković 2026a).[12]

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.

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

4.3.3. 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.)

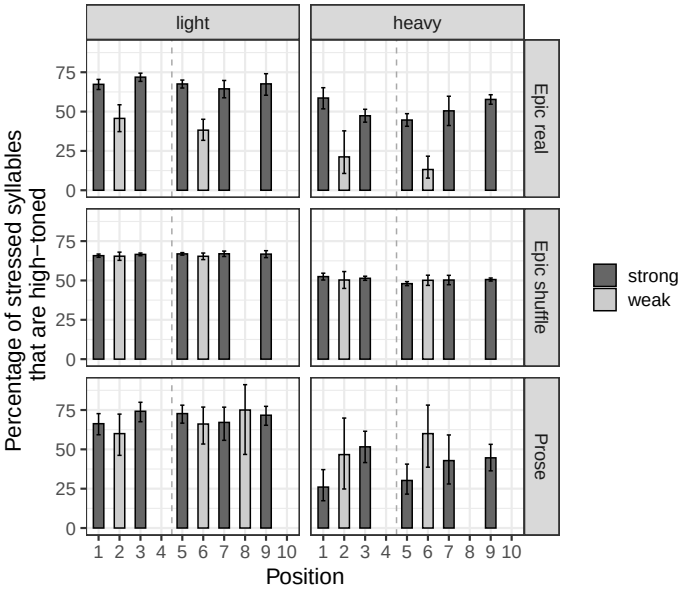
4.4. Section conclusion

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 be 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.

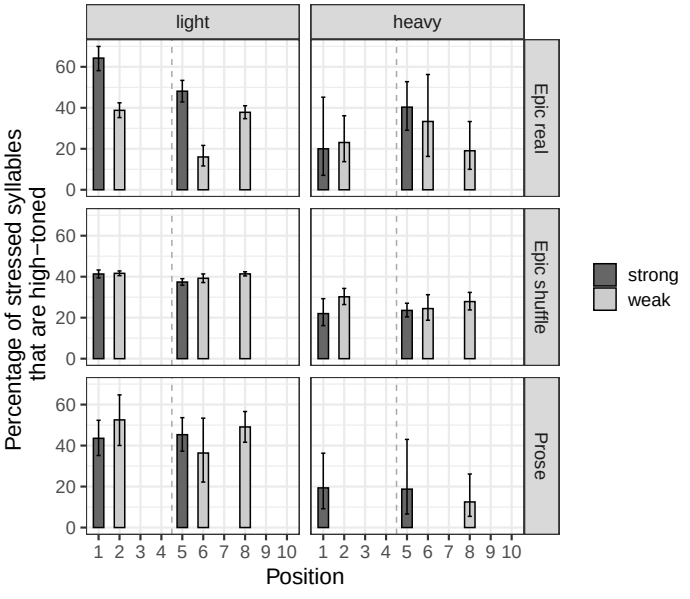
The inferential statistics support the following two main generalizations. First, the meter is sensitive to weight in stressed syllables. While stressed syllables show a particularly strong tendency to be heavy in position 9 (the final strong), weight is also regulated elsewhere in the line. Second, 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



(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.

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 (13) (Hanson and Kiparsky 1996).

(13) 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 (13a), 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 (13b), 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 (14).

(14) STRONG⇒HEAVY
Assign a violation for every strong position that contains a light 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 (15; e.g. Blumenfeld 2016, cf. Hayes et al. 2012).

(15) STRESS⇒STRONG
Assign a violation for every stressed syllable assigned to a weak position.

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, 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 (§4.3.2).

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 interact only 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, Ryan 2026).[13] The BCMS epic decasyllable therefore constitutes the first documented case of tone regulation in Indo-European.

[13] Lubotsky (1995) argues that the Sanskrit Rig-Veda exhibits a tendency for accentual responsion between lines of a couplet. Even if that were the case (cf. Ryan 2026, 18–20 for a statistical reassessment), Lubotsky (1995) explicitly agrees with the traditional view that the accent plays no systematic role in Vedic meter; responsion is a type of local parallelism.

Moving to (13b) (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 (§4.3.2), 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 (OT; Prince and Smolensky 2004).[14] Arguments for HG over OT in the context of variation are provided by Zuraw and Hayes (2017), Smith and Pater (2020), Hayes (2022), among others.

A candidate's Harmony score $\mathscr{H}_i$ is defined as the negative weighted sum of its violations (16), where w_k is the weight of constraint k (always nonnegative; Keller 2000, Potts et al. 2010), 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).

(16) Harmony

$$\mathscr{H}_i = -\sum_{k \in C} w_k c_{ik}$$

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 $e^{\mathscr{H}_i}$ divided by the sum of exponentiated Harmonies of all candidates j in the candidate set X (17).

(17) Candidate probability in MaxEnt

$$p_i = \frac{e^{\mathscr{H}_i}}{\sum_{j \in X} e^{\mathscr{H}_j}}$$

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 (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).

[14] For an accessible introduction to OT and constraint-based phonology more generally, see McCarthy (2008). For MaxEnt specifically, see Hayes and Moore-Cantwell (2025).

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's (2008) phonotactic model. 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.[15]

In (18–19), we present constraints that govern the independent mapping of stress, weight, and tone.

(18) a. Stress⇒Strong
Assign a violation for every stressed syllable assigned to a weak position.

b. Heavy⇒Strong
Assign a violation for every heavy syllable assigned to a weak position.

c. High⇒Strong
Assign a violation for every high-toned syllable assigned to a weak position.

(19) a. Strong⇒Stress
Assign a violation for every strong position containing an unstressed syllable.

b. Strong⇒Heavy
Assign a violation for every strong position containing a light syllable.

c. Strong⇒High
Assign a violation for every strong position not containing a high-toned syllable.

To model stress–weight interaction, we adopt the constraints in (20) (Ryan 2017).

(20) a. Stress⇒(Strong⇒Heavy)
Assign a violation for every stressed light syllable assigned to a strong position.

b. Stress⇒(Heavy⇒Strong)
Assign a violation for every stressed heavy syllable assigned to a weak position.

The tableau in (21) shows constraints (20a–20b) at work.

(21)

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

The tonal counterparts of the constraints in (20) are defined in (22). Their effect is illustrated in tableau (23).

[15] These constraints could equivalently be expressed as conjunctions, since $A \rightarrow B \equiv \neg(A \wedge \neg B)$. For example, constraint (18a) could be expressed as *Stress/Weak, and (19a) could be expressed as *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.).

(22) a. STRESS⇒(STRONG⇒HIGH)
Assign a violation for every stressed non-high-toned syllable assigned to a strong position.

b. STRESS⇒(HIGH⇒STRONG)
Assign a violation for every stressed high-toned syllable assigned to a weak position.

(23)

	STRESS⇒(STRONG⇒HIGH)	STRESS⇒(HIGH⇒STRONG)
a. ˈσ_{μ}/S	*	
b. ˈ$\acute{\sigma}_{\mu}$/W		*

The constraint system adopted here 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% of tokens). By contrast, stressed toneless lights, with just one degree of prominence, are the stressed syllable type most likely to be misaligned (43% of tokens misaligned); see (24).

(24)

	STRESS⇒STRONG	HEAVY⇒STRONG	HIGH⇒STRONG	STRESS⇒(HEAVY⇒STRONG)	STRESS⇒(HIGH⇒STRONG)	misalignment rate	misaligned tokens	total
a. ˈ$\acute{\sigma}_{\mu\mu}$/W	*	*	*	*	*	9%	167	1,880
b. ˈ$\sigma_{\mu\mu}$/W	*	*		*		19%	371	1,923
c. ˈ$\acute{\sigma}_{\mu}$/W	*		*		*	19%	851	4,478
d. ˈσ_{μ}/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 (25), following the demonstration that weight mapping is stricter at the end of lines (Maretić 1901, Jakobson 1966, Zec 2008). Constraint (25a) requires that stressed syllables in the ninth position be heavy (§4.3.2). Furthermore, constraints (25b–25c) are the cadential counterparts of (18b–19b): (25b) penalizes heavy syllables in the tenth position, while (25c) militates against the ninth positions that do not contain a heavy syllable.

(25) a. STRESS⇒(STRONG⇒HEAVY)$_{\text{cadence}}$
Assign a violation for every stressed light syllable assigned to a strong position in the cadence.

b. HEAVY⇒STRONG$_{\text{cadence}}$
Assign a violation for every heavy syllable assigned to a weak position in the cadence.

c. STRONG⇒HEAVY$_{\text{cadence}}$
Assign a violation for every strong position in the cadence containing a light syllable.

In sum, there are thirteen constraints in the initial set (18–25), which will next be subjected to a selection procedure.

6.1.2. Model selection

Using a fixed random seed for reproducibility, we randomly partitioned 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 HOLDOUT (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. The criterion for weight search is maximum likelihood (Myung 2003). The `maxent.ot` package uses the limited-memory BFGS algorithm to find the weights that maximize the likelihood of the observed data given the grammar. In this subsection, we report the selection 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 affect constraint selection or weight estimation, which prevents information leakage and provides an unbiased estimate of model generalization.

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 of candidate constraints 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 (19) were not included in the final model, confirming that weak positions are the primary locus of regulation in the epic decasyllable (Zec 2008; §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.[16] 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.

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.

[16] While the p-value of STRESS⇒(HEAVY⇒STRONG) is borderline here, see §6.2 for further support.

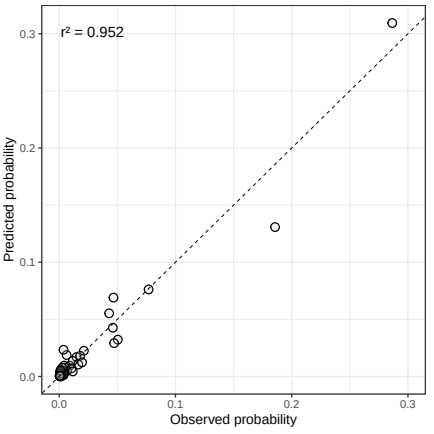


Figure 8.: Scatterplot of observed vs. predicted frequencies of line types in the training dataset.

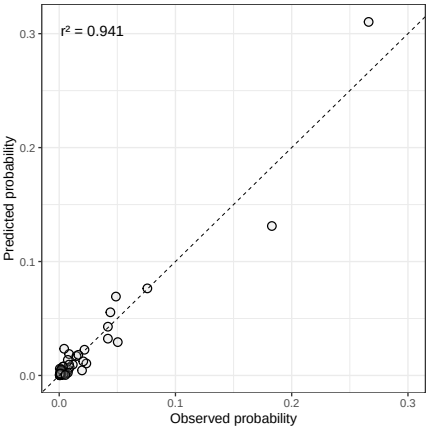


Figure 9.: Model fit to test data.

As a second test, we evaluate the model on line types argued by previous scholarship to be unmetrical. A metrical grammar should not only assign probability to the attested line types; it should also effectively 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 in the epic decasyllable. Consider the lines in (26), which instantiate structures that violate Jakobson (1966)'s metrical constants, structural properties shared by all verse lines in this oral tradition. All lines in (26) 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, at least one word boundary must occur before an odd position (Jakobson 1966, *ibid*.). All lines in (26) violate this principle. As shown, our model assigns extremely high penalties to these lines, correctly driving their predicted probabilities towards zero.[17]

(26) a. pa ˈkréɛ.nuu ˈtád ‖ da ˈbrá.ta ˈnáa.ʤɛɛɛ ˈɔ́ɔn
so moved then that brother find he
'So he moved then to find (his) brother.' ($\mathscr{H}$ = −13.58)

[17] Lines (26a–26b) are artificially created to violate Jakobson's constants, while (26c) is a permutation of a real line produced by the legendary guslar Avdo Međedović (*The Wedding of Vlahinjić Alija*, sung version, 2,486; Parry 1980). The attested counterpart of (26c) (ak ˈnɛ=dɔ́ɔ.ʤɛɛʃ ‖ da ˈbɔ́ɔg ˈléɛ.pii ˈʧúu.ʋaa) is less marked but still metrically complex ($\mathscr{H}$ = −6.60) due to a long sequence of heavy syllables. The poet chose the word order that maps these heavies in the least costly fashion.

b. kad ˈtɔ́ɔ ˈʧúu ˈɔ́ɔn ‖ ˈtád ˈsíi.ʤɛɛ ˈɔ́.pɛɛt ˈtúu
when that heard he then descends again there
‘When he heard that, he descended there again.’ ($\mathscr{H}$ = −15.56)

c. ak ˈnɛ=dɔ́ɔ.ʤɛɛʃ ‖ da ˈʧúu.ʋaa ˈlɛ́ɛ.pii ˈbɔ́ɔg
if not=come that saves handsome God
‘If you don’t come, may dear God save us!’ ($\mathscr{H}$ = −11.64)

Our model captures the effects of Jakobson’s (1966) metrical constants despite the fact that these are not directly encoded in the model. Violations of metrical constants induce costly misalignment of multiply-prominent syllables (e.g. stressed high-toned syllables or stressed heavies), which is penalized by constraints on stress-tone and stress-weight combinations.

6.1.4. Gradient metricality, cumulativity, and harmonic bounding

In lieu of a local conclusion, we situate our findings in recent theoretical work and discuss why MaxEnt is a sensible choice in the present modeling setup. 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.[18] 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–d), which have relatively low complexity (i.e. markedness; Halle and Keyser 1971), still have triple-digit frequency. Mid-complexity lines (roughly e–g) are less frequent, and the high-complexity candidates (h–j) occur only a handful of times (i–j are hapaxes) and are assigned vanishing probabilities by the model (see Figure 8).

As evident from Table 7, constraint violations interact cumulatively (Jäger and Rosenbach 2006): Line-type frequency drops the more constraints lines violate, and the more violations of individual constraints they incur. The data thus display both GANGING-UP CUMULATIVITY, whereby coincident violations of two or more constraints jointly reduce the production frequency of verse lines (Smith and Pater 2020, Breiss and Albright 2022), and COUNTING CUMULATIVITY, whereby line types get incrementally less frequent with multiple violations of a single constraint (Kawahara and Kumagai 2021, Kim 2022).

Ganging-up cumulativity manifests itself in the fact that coincident violations of multiple independent constraints are rarer than individual violations in our corpus. Consider example (f) in Table 7, where the violations of HEAVY⇒STRONG and STRESS⇒(STRONG⇒HEAVY)$_{\text{cadence}}$ add up to reduce the probability of the doubly-violating line type to 1%, compared to 18% and 2% for the singly-violating lines in (b) and (e), respectively. Examples (a), (b), and (d) illustrate counting cumulativity. These line types fare identically on all metrical constraints except for HEAVY⇒STRONG: (a) incurs none, (b) incurs one, and (d) incurs two HEAVY⇒STRONG violations. Their observed probability in the corpus decreases monotonically with the violation count, from 28% for the violation-free line type, to 18% for the line type incurring a single violation, and only 4% for the line type incurring two violations. For other cases where the two types of cumulativity coexist in the grammar, see Jäger and Rosenbach (2006), McPherson (2016), Breiss (2020), Kawahara and Breiss (2021).

In MaxEnt and HG more generally, interaction terms between constraints, known as constraint conjunctions (Smolensky 2006), are sometimes invoked to model extreme cases of superlinear cumulativity (Green and Davis 2014, Shih 2017, Milenković 2026b), whereby the coincident violation of multiple constraints yields a steeper probability decline than the combination of their independent effects. We tested the utility of several conjoined constraints and found that none significantly improve model fit (see Appendix B for more detail). This suggests that our current model captures the full range of ganging-up interactions in the data with no appeal to interaction terms.

The identification of pervasive ganging-up and counting cumulativity in the BCMS folk meter has theoretical implications, strongly supporting HG over strict-ranking OT. All variants of HG predict

[18] Line type frequencies in Table 7 are reported for the entire corpus, not just the training set.

	Example line	Heavy⇒Strong 0.86	Stress⇒(Heavy⇒Strong) 1.12	Stress⇒(Strong⇒Heavy)$_{cad}$ 2.62	Heavy⇒Strong$_{cad}$ 0.64	Stress⇒(High⇒Strong) 1.40	$\mathscr{H}$	*N*
a.	ˈtɔ́ɔ jɛ ˈmáar.kɔ ‖ ˈpɔ.slú.ʃa.ɔ ˈmáaj.ku ‘In that Marko listened to his mother.’ (II, 72:14)						0.00	1,042
b.	ˈní.jɛ ˈvɛ́.ʨɛɛɛ ‖ ˈʧú.dɔ ˈna.stá.nu.lɔ ‘No greater wonder had ever occurred.’ (II, 39:2)	*					−0.86	695
c.	ˈtí.xɔ ˈjá.ʃtɛ ‖ ˈgɔ́.spɔ.dɔ ˈsʋá.tɔ.ʋi ‘Ride quietly, wedding party.’ (IV, 13:31)					*	−1.40	288
d.	ˈjɛ́d.naak ˈvɛ.zíir ‖ ˈtsá.ra ˈpɔ.slú.ʃa.ɔ ‘The vizier immediately obeyed the tsar.’ (III, 11:58)	**					−1.72	164
e.	ˈnɛ bí l ˈtsá.ra ‖ u ʒi.ˈvɔ.tú ˈna.ʃlí ‘In the hope of finding the tsar alive.’ (II, 33:14)			*			−2.62	81
f.	ˈu.drí ˈbrá.tɛ ‖ a.kɔ ˈbɔ́.ʌɛɛ ˈmɔ́.ʒɛʃ ‘Strike, brother, if you can do better.’ (II, 30:260)	*		*			−3.48	39
g.	ˈnút ˈdí.ʋɛ.rɛ ‖ ˈzláat.nii ˈmɔ́ɔj ˈpɻ́.stɛ.nɛ ‘Here, my brother-in-law, my golden ring.’ (I, 342:89)	*				**	−3.66	16
h.	ˈtɛ́r ˈpri.síi.tsaa ‖ ˈdʋáa ˈzláat.na ˈgaaj.tá.na ‘Then she cut the two golden cords.’ (I, 342:65)	***	**			*	−6.22	2
i.	ˈɛ́.tɔ ˈʨii.rɛ́ɛ ‖ a ˈnɛ́ɛ.maa ˈsʋá.tɔɔ.ʋaa ‘Here comes Ćira, but no wedding party.’ (III, 76:130)	***	*		*	**	−7.14	1
j.	da ˈjáa ˈví.diim ‖ ˈʃtá ˈí.maa u ˈgɔ.rí ‘So that I may see what is in the hills.’ (III, 47:209)	**	*	*		**	−8.26	1

Table 7.: Correlation between Harmony and line type frequency.

ubiquitous constraint cumulativity because in HG, all constraint violations incurred by a candidate count towards its Harmony. By contrast, OT employs “fast-and-frugal” evaluation (Kawahara and Breiss 2021): The winner is determined by the top-ranking constraint(s), such that violations of lower-ranked constraints play no role when their higher-ranked counterparts discriminate between candidates. As a result, Classical OT categorically rules out both ganging-up and counting cumulativity, at least without nondefault machinery (Pater 2009, 2016, Potts et al. 2010). Stochastic OT (Boersma and Hayes 2001), a widely used variant of OT that accommodates variable output, shows only limited success in modeling ganging-up cumulativity (Jäger and Rosenbach 2006, Hayes and Londe 2006, Maslova 2007, Jarosz

2010), which makes it an overly restrictive theory of this phenomenon (Smith and Pater 2020, Flemming and Magri 2026). Moreover, like Classical OT, stochastic OT predicts no counting cumulativity (Jäger and Rosenbach 2006, Keller 2006, Kawahara 2020).

Another key property of MaxEnt that drives its success in modeling gradient metricality is the absence of strict harmonic bounding. Two candidates are in a simple harmonic bounding relation when one (the bounder) incurs a proper subset of the violations incurred by the other (the bounded) (Prince and Smolensky 2004, Samek-Lodovici and Prince 2002), as in (27). Unlike most other constraint-based frameworks, MaxEnt can generate harmonically bounded candidates with nonnegligible probability, as in (28). Despite this flexibility vis-à-vis harmonic bounding, recall that MaxEnt still handles near-categorical exclusion by assigning poorly performing candidates probabilities that are extremely close to zero (see Figures 8–9).

(27) Simple harmonic bounding

	$\mathbb{C}_1$	$\mathbb{C}_2$	$\mathbb{C}_n$
a. bounder		*	
b. bounded		*	*!

(28) MaxEnt assigns nonzero probability to the bounded

	$\mathbb{C}_1$ 1	$\mathscr{H}$	$e^{\mathscr{H}}$	p
a. bounder		0.00	1.00	0.73
b. bounded	−1	−1.00	0.37	0.27

In our corpus, the violation-free line type simply harmonically bounds every line type that incurs at least one violation, that is, every other line type. The MaxEnt model distributes probability over all line types, proportional to their Harmony. Frameworks that uphold strict harmonic bounding cannot model gradient metricality in the present setup, given the harmonic bounding relations between attested line types. Stochastic OT is a case in point: Although the framework generally accommodates gradience and variation, it fails to capture gradient metricality, since it observes strict harmonic bounding. Indeed, the stochastic OT model of our corpus data[19] assigns all probability to the violation-free line type; all other, harmonically bounded, line types receive zero probability irrespective of their metrical well-formedness. In the context of gradient metricality and acceptability judgments, MaxEnt's ability to assign nonzero probability to simply harmonically bounded forms is therefore a virtue, not a liability (Hayes and Wilson 2008, Hayes and Moore-Cantwell 2011, Hayes et al. 2012, Hayes and Schuh 2019, Kumagai 2023).

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;

[19]The stochastic OT model was run using OT-Soft (Hayes et al. 2013). The procedure uses the Gradual Learning Algorithm (Boersma 1998), with Magri's (2012) update rule, which addresses the convergence issues identified by Pater (2008). Otherwise, default settings were used: The grammar is trained over 999 learning trials and tested over 1,000,000 evaluation trials; starting and final plasticity are set to 2.00 and 0.01, respectively; all constraints have an initial ranking value of 100; and evaluation noise is set to 2.

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 in 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's (2022) 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.[20] 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. A constraint contributes to meter beyond its baseline effect when reducing it to its prose weight leads to a BIC increase of at least 2. The results are shown in Table 8.[21]

Constraint	weight (epic)	weight (prose)	weight difference (epic − prose)	ΔBIC
Heavy⇒Strong	0.86	0.76	0.10	15.79
Stress⇒(Heavy⇒Strong)	1.07	0.18	0.89	390.99
Stress⇒(Strong⇒Heavy)$_{\text{cadence}}$	2.59	0.77	1.82	1,131.40
Heavy⇒Strong$_{\text{cadence}}$	0.57	0.14	0.43	81.10
Stress⇒(High⇒Strong)	1.38	0.84	0.54	288.76

Table 8.: Meter-specific constraint contribution: Reducing each constraint to its prose baseline effect results in a considerably higher BIC (i.e. worse model fit).

As Table 8 shows, all five constraints receive more weight in the epic than in the 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 (§4.3.2). Stress⇒(High⇒Strong) 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 (§4.3.3). The BIC comparisons show that suppressing any constraint to its baseline contribution leads to a significant BIC increase, signaling poorer model fit.[22] This confirms that each constraint has an independent metrical contribution.

[20] 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 $\mu = 0$.

[21] 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 the training set, which comprises one half of the metrical corpus, whereas the latter was fit to the full corpus.

[22] A BIC increase of at least 2 provides positive evidence for retaining the omitted (here, reduced) constraint, whereas ΔBIC ≥ 10 constitutes very strong evidence for retention (Raftery 1995).

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.

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's (2017) 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. One property of MaxEnt that makes it a particularly well-suited model of gradient metricality is that it can assign nonzero probability to harmonically bounded line types. This property is an advantage given that attested verse lines are often harmonically bounded by their rivals (Hayes and Moore-Cantwell 2011, Hayes et al. 2012, Hayes and Schuh 2019). In fact, all line types in our corpus other than the perfectly metrical, violation-free type are harmonically bounded by at least one other candidate.

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).

Supplementary material. To view supplementary material for this article, please visit http://doi.org/10.1017/Sxxx850725xxxx

Data availability. Code for all simulations and experiments is available at https://github.com/morph-order

Competing interests. None.

Ethical standard. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

References

Allen, W. S. (1973). *Accent and rhythm*. Cambridge University Press, Cambridge.

Arlot, S. and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4:40–79.

Arnold, E. V. (1905). *Vedic metre in its historical development*. University Press, Cambridge.

Bailey, J. (1975). *Toward a statistical analysis of English verse. The iambic tetrameter of ten poets*. Peter de Ridder, Lisse.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Batinić, P. (1975). *The meter of the Serbo-Croatian oral epic decasyllable*. PhD thesis, Columbia University.

Bickmore, L. S. (1995). Tone and stress in Lamba. *Phonology*, 12(3):307–341.

Blumenfeld, L. (2016). Generative metrics: An overview. *Language and Linguistics Compass*, 10(9):413–430.

Boersma, P. (1998). *Functional Phonology: Formalizing the interactions between articulatory and perceptual drives*. PhD thesis, University of Amsterdam.

Boersma, P. and Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, 32(1):45–86.

Breiss, C. (2020). Constraint cumulativity in phonotactics: Evidence from artificial grammar learning studies. *Phonology*, 37(4):551–576.

Breiss, C. and Albright, A. (2022). Cumulative markedness effects and (non-) linearity in phonotactics. *Glossa: a journal of general linguistics*, 7(1):1–32.

de Lacy, P. (2002). The interaction of tone and stress in Optimality Theory. *Phonology*, 19(1):1–32.

deCastro Arrazola, V. (2018). Testing the robustness of final strictness in verse lines. *Studia Metrica et Poetica*, 5(2):55–76.

Della Pietra, S., Della Pietra, V., and Lafferty, J. (1997). Inducing features of random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(4):380–393.

Devroye, L. and Wagner, T. (1979). Distribution-free performance bounds for potential function rules. *IEEE Transactions on Information Theory*, 25(5):601–604.

Fabb, N. and Halle, M. (2008). *Meter in poetry: A new theory*. Cambridge University Press, Cambridge.

Farris-Trimble, A. W. (2008). *Cumulative faithfulness effects in phonology*. PhD thesis, Indiana University.

Flemming, E. (2021). Comparing MaxEnt and Noisy Harmonic Grammar. *Glossa: a journal of general linguistics*, 6(1):1–42.

Flemming, E. and Magri, G. (2026). Gang effects: The perspective from variation. In Sande, H., Asmah, R., Beguš, G., Dabkowski, M., Garrett, A., Kavitskaya, D., Russell, K., and Yu, A., editors, *Proceedings of the 2025 Annual Meeting on Phonology*, Amherst, MA. University of Massachusetts Amherst Libraries.

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

Gasparov, M. L. (1980). Quantitative methods in Russian metrics: Achievements and prospects. *Russian poetics in translation*, 7:1–19.

Gasparov, M. L. (1987). Probability model of verse (English, Latin, French, Italian, Spanish, Portuguese). *Style*, 21(3):322–358. Translated by Marina Tarlinskaja.

Goldsmith, J. (1987). Tone and accent, and getting the two together. In *Annual Meeting of the Berkeley Linguistics Society*, volume 13, pages 88–104.

Goldwater, S. and Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In Spenader, J., Eriksson, A., and Dahl, Ö., editors, *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, pages 111–120, Stockholm. Stockholm University, Department of Linguistics.

Golston, C. and Riad, T. (2000). The phonology of Classical Greek meter. *Linguistics*, 38(1):99–167.

Gordon, M. (2006). *Syllable weight: phonetics, phonology, typology*. Routledge, New York.

Gordon, M. (2023). The phonetic basis for tone-stress interactions: A cross-linguistic study. In van de Weijer, J., editor, *Representing Phonological Detail, Part II: Syllable, Stress, and Sign*, pages 171–190. Walter de Gruyter, Berlin.

Green, C. and Davis, S. (2014). Superadditivity and limitations on syllable complexity in Bambara words. In Farris-Trimble, A. and Barlow, J., editors, *Perspectives on phonological theory and development, in honor of Daniel A. Dinnsen*, pages 223–247. John Benjamins, Amsterdam.

Gunkel, D. C. and Ryan, K. (2011). Hiatus avoidance and metrification in the Rigveda. In Jamison, S. W., Melchert, H. C., and Vine, B., editors, *Proceedings of the 22nd Annual UCLA Indo-European Conference*, pages 53–68, Bremen. Hempen.

Hall, D. C. (2006). Modeling the linguistics-poetics interface. In Dresher, B. E. and Friedberg, N., editors, *Formal approaches to poetry: Recent developments in metrics*, pages 233–49. Mouton de Gruyter, Berlin.

Halle, M. and Keyser, S. J. (1966). Chaucer and the study of prosody. *College English*, 28(3):187–219.

Halle, M. and Keyser, S. J. (1971). Theory of prosody, continued: Illustration and defense of a theory of the iambic pentameter. *College English*, 33(2):154–176.

Hanson, K. and Kiparsky, P. (1996). A parametric theory of poetic meter. *Language*, 72:287–335.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning*. Springer, New York.

Hayes, B. (1983). A grid-based theory of English meter. *Linguistic Inquiry*, 14(3):357–393.

Hayes, B. (1988). Metrics and phonological theory. In Newmeyer, F. J., editor, *Linguistics: The Cambridge Survey: Volume 2, Linguistic Theory: Extensions and Implications*, pages 220–249. Cambridge University Press, Cambridge.

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

Hayes, B. (2022). Deriving the wug-shaped curve: A criterion for assessing formal theories of linguistic variation. *Annual Review of Linguistics*, 8(1):473–494.

Hayes, B. (2023). A brief introduction to meter. Talk at M@100, MIT.

Hayes, B. and Londe, Z. C. (2006). Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology*, 23(1):59–104.

Hayes, B. and Minkova, D. (2023). A MaxEnt analysis of the meter of Beowulf. Poster at M@100, MIT.

Hayes, B. and Moore-Cantwell, C. (2011). Gerard Manley Hopkins' sprung rhythm: Corpus study and stochastic grammar. *Phonology*, 28(2):235–282.

Hayes, B. and Moore-Cantwell, C. (2025). A guide to analysis in MaxEnt Optimality Theory. Ms., UCLA. February 2025 version.

Hayes, B. and Schuh, R. G. (2019). Metrical structure and sung rhythm of the Hausa rajaz. *Language*, 95(2):e253–e299.

Hayes, B., Tesar, B., and Zuraw, K. (2013). *OT-Soft 2.5*. Software package.

Hayes, B. and White, J. (2015). Saltation and the p-map. *Phonology*, 32(2):267–302.

Hayes, B. and Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39(3):379–440.

Hayes, B., Wilson, C., and Shisko, A. (2012). Maxent grammars for the metrics of Shakespeare and Milton. *Language*, 88(4):691–731.

Henriksson, E. (2022). *Greek meter: an approach using metrical grids and Maxent*. PhD thesis, Helsingin yliopisto.

Hyman, L. M. (2006). Word-prosodic typology. *Phonology*, 23(2):225–257.

Inkelas, S. and Zec, D. (1988). Serbo-Croatian pitch accent: the interaction of tone, stress, and intonation. *Language*, 64:227–248.

Ivić, P. (1958). *Die serbokroatischen Dialekte. Ihre Struktur und Entwicklung*. Mouton, The Hague.

Jäger, G. and Rosenbach, A. (2006). The winner takes it all—almost: Cumulativity in grammatical variation. *Linguistics*, 44:937–971.

Jakobson, R. (1966). Slavic epic verse: Studies in comparative metrics. In *Selected writings*, volume 4, pages 414–463. Mouton, The Hague & Paris.

Janson, T. (1975). *Prose rhythm in Medieval Latin from the 9th to the 13th century*. Almqvist & Wiksell International, Stockholm.

Jarosz, G. (2010). Implicational markedness and frequency in constraint-based computational models of phonological learning. *Journal of Child Language*, 37(3):565–606.

Kager, R. (1989). *A metrical theory of stress and destressing in English and Dutch*. PhD thesis, Utrecht University.

Karadžić, V. S. (1824). Predgovor. In *Narodne srpske pjesme, skupio ih i na svijet izdao Vuk Stef. Karadžić. Knjiga prva, u kojoj su različne ženske pjesme*. Breitkopf & Härtel, Leipzig.

Karadžić, V. S. (1841–1862). *Srpske narodne pjesme. Knjiga prva, u kojoj su različne 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)*. U štampariji jermenskoga manastira, Vienna.

Karadžić, V. S. (1852). *Srpski rječnik istumačen njemačkijem i latinskijem riječima*. Vienna.

Kawahara, S. (2020). A wug-shaped curve in sound symbolism: The case of Japanese Pokémon names. *Phonology*, 37(3):383–418.

Kawahara, S. and Breiss, C. (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(1):1–29.

Kawahara, S. and Kumagai, G. (2021). What voiced obstruents symbolically represent in Japanese: Evidence from the Pokémon universe. *Journal of Japanese Linguistics*, 37(1):3–24.

Keller, F. (2000). *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. PhD thesis, The University of Edinburgh.

Keller, F. (2006). Linear Optimality Theory as a model of gradience in grammar. In Fanselow, G., Féry, C., Schlesewsky, M., and Vogel, R., editors, *Gradience in Grammar: Generative Perspectives*, pages 270–288. Oxford University Press, Oxford.

Kim, S. (2022). A MaxEnt learner for super-additive counting cumulativity. *Glossa: a journal of general linguistics*, 7(1):1–34.

Kiparsky, P. (1968). Metrics and morphonemics in the Kalevala. In Gribble, C. E., editor, *Studies presented to Roman Jakobson by his students*, pages 137–148. Slavica Publishers, Cambridge, MA.

Kiparsky, P. (1975). Stress, syntax, and meter. *Language*, 51(3):576–616.

Kiparsky, P. (1977). The rhythmic structure of English verse. *Linguistic Inquiry*, 8(2):189–247.

Kiparsky, P. (2020). Metered verse. *Annual Review of Linguistics*, 6(1):25–44.

Kumagai, G. (2023). Express [p] in expressive phonology: Analysis of a nicknaming pattern using 'princess' in Japanese. *Phonology*, 40(3–4):267–290.

Legendre, G., Miyata, Y., and Smolensky, P. (1990). Harmonic grammar—a formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. In *Proceedings of the twelfth annual conference of the Cognitive Science Society*, pages 388–395, Mahwah, NJ. Lawrence Erlbaum Associates.

Lehiste, I. and Ivić, P. (1986). *Word and sentence prosody in Serbocroatian*. MIT Press, Cambridge, MA.

Lord, A. B. (1960). *The singer of tales*. Harvard University Press, Cambridge, MA.

Lubotsky, A. (1995). Accentuation in the technique of the Vedic poets. In Semeka-Pankratov, E., editor, *Studies in poetics: commemorative volume Krystyna Pomorska*, pages 515–534. Slavica Publishers, Columbus, Ohio.

Maas, P. (1962). *Greek metre*. Clarendon Press, Oxford.

Macdonell, A. A. (1910). *Vedic grammar*. Verlag von Karl J. Trübner, Strassburg.

Magri, G. (2012). Convergence of error-driven ranking algorithms. *Phonology*, 29(2):213–269.

Maretić, T. (1901). O građi našega narodnog deseterca. *Nastavni vjesnik*, 10(2):145–175.

Maretić, T. (1907). Metrika narodnih naših pjesama. *Rad Jugoslavenske akademije znanosti i umjetnosti*, 168:1–112.

Maslova, E. (2007). Stochastic OT as a model of constraint interaction. In Zaenen, A., Simpson, J., King, T. H., Grimshaw, J., Maling, J., and Manning, C., editors, *Architectures, Rules, and Preferences: Variations on Themes by Joan W. Bresnan*, pages 513–528. CSLI Publications, Stanford.

Matić, S. (1964). *Naš narodni ep i naš stih.* Matica srpska, Novi Sad.

Mayer, C., Tan, A., and Zuraw, K. R. (2024). Introducing `maxent.ot`: an R package for Maximum Entropy constraint grammars. *Phonological Data and Analysis*, 6(4):1–44.

McCarthy, J. J. (2008). *Doing Optimality Theory: Applying theory to data*. Blackwell, Malden, MA.

McPherson, L. (2016). Cumulativity and ganging in the tonology of Awa suffixes. *Language*, 92(1):e38–e66.

McPherson, L. and Hayes, B. (2016). Relating application frequency to morphological structure: the case of Tommo So vowel harmony. *Phonology*, 33(1):125–167.

McPherson, L. and Ryan, K. M. (2018). Tone-tune association in Tommo So (Dogon) folk songs. *Language*, 94(1):119–156.

Milenković, A. (2026a). *Metrical grammar: Variation, gradience, and cumulativity at the syntax-phonology interface*. PhD thesis, Harvard University.

Milenković, A. (2026b). Superadditive cumulativity in categorical prosodic patterns: Prosodic minimality in Bosnian/Croatian/Montenegrin/Serbian. *Phonology*, 43(e4):1–33.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, 47(1):90–100.

Nikolić, B. (1970). *Osnovi mlađe novoštokavske akcentuacije*. Institut za srpskohrvatski jezik, Beograd.

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

Pater, J. (2008). Gradual learning and convergence. *Linguistic Inquiry*, 39(2):334–345.

Pater, J. (2009). Weighted constraints in generative linguistics. *Cognitive Science*, 33(6):999–1035.

Pater, J. (2016). Universal grammar with weighted constraints. In Pater, J. and McCarthy, J., editors, *Harmonic grammar and harmonic serialism*, pages 1–46. London: Equinox.

Potts, C., Pater, J., Jesney, K., Bhatt, R., and Becker, M. (2010). Harmonic grammar with linear programming: from linear systems to linguistic typology. *Phonology*, 27(1):77–117.

Prince, A. (1990). Quantitative consequences of rhythmic organization. In Ziolkowski, M., Noske, M., and Deaton, K., editors, *Papers from the 26th Regional Meeting of the Chicago Linguistic Society: Volume 2: The Parasession*, pages 355–398, Chicago. Chicago Linguistic Society.

Prince, A. and Smolensky, P. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell, Malden, MA. Technical Report, Rutgers University and University of Colorado at Boulder, 1993. Revised version Blackwell, 2004.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25:111–163.

Raven, D. S. (1962). *Greek metre: an introduction*. Faber & Faber, London.

Ružić, Ž. (1975). *Srpski jamb i narodna metrika*. Institut za književnost i umetnost, Beograd.

Ryan, K. M. (2011). Gradient syllable weight and weight universals in quantitative metrics. *Phonology*, 28(3):413–454.

Ryan, K. M. (2016). Phonological weight. *Language and Linguistics Compass*, 10(12):720–733.

Ryan, K. M. (2017). The stress–weight interface in metre. *Phonology*, 34(3):581–613.

Ryan, K. M. (2019). *Prosodic Weight: Categories and continua*. Oxford University Press, Oxford.

Ryan, K. M. (2026). The role of pitch accent in Vedic Sanskrit poetics. *Phonological Data and Analysis*, 8(e5):1–26.

Samek-Lodovici, V. and Prince, A. (2002). Fundamental properties of harmonic bounding. Technical Report RuCCS-TR-71, Rutgers Center for Cognitive Science. Revised as ROA-785 (2005).

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Selkirk, E. (1996). The prosodic structure of function words. In Morgan, J. L. and Demuth, K., editors, *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, pages 187–213. Erlbaum, Mahwah, NJ.

Shih, S. S. (2017). Constraint conjunction in weighted probabilistic grammar. *Phonology*, 34(2):243–268.

Smith, B. W. and Pater, J. (2020). French schwa and gradient cumulativity. *Glossa: a journal of general linguistics*, 5(1):1–33.

Smith, L. and Anttila, A. T. (2025). Metrical tension and prose cadence. In *Proceedings of the Annual Meetings on Phonology*, volume 1. University of Massachusetts Amherst Libraries.

Smolensky, P. (2006). Optimality in phonology II: Harmonic completeness, local constraint conjunction, and feature domain markedness. In Smolensky, P. and Legendre, G., editors, *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar. Vol. 2: Linguistic and philosophical implications*, pages 27–160. MIT Press, Cambridge, MA.

Tarlinskaja, M. (1976). *English verse: Theory and history*. Mouton, The Hague and Paris.

Tarlinskaja, M. G. and Teterina, L. M. (1974). Verse–prose–metre. *Linguistics*, 12:63–86.

Vaillant, A. (1932). Les Chants épiques des Slaves du Sud. *Revue des cours et conférences*, 4:309–326.

Werle, A. (2009). *Word, phrase, and clitic prosody in Bosnian, Serbian, and Croatian*. PhD thesis, University of Massachusetts Amherst.

West, M. L. (1987). *Introduction to Greek metre*. Clarendon Press, Oxford.

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, 30(5):945–982.

Yip, M. (2002). *Tone*. Cambridge University Press, Cambridge.

Zec, D. (1999). Footed tones and tonal feet: rhythmic constituency in a pitch-accent language. *Phonology*, 16(2):225–264.

Zec, D. (2000). O strukturi sloga u srpskom jeziku. *Južnoslovenski filolog*, 56:435–448.
Zec, D. (2005). Prosodic differences among function words. *Phonology*, 22(1):77–112.
Zec, D. (2008). The prosodic word as a unit in poetic meter. In Hanson, K. and Inkelas, S., editors, *The Nature of the Word: Essays in Honor of Paul Kiparsky*, pages 63–94. MIT Press, Cambridge, MA.
Zec, D. and Zsiga, E. (2010). Interaction of tone and stress in standard serbian: Phonological and phonetic evidence. In Browne, W., Cooper, A., Fisher, A., Kesici, E., Predolac, N., and Zec, D., editors, *Formal approaches to Slavic linguistics 18: The Cornell Meeting 2008*, pages 536–555. Michigan Slavic Publications.
Zec, D. and Zsiga, E. (2022). Tone and stress as agents of cross-dialectal variation: The case of Serbian. In Kubozono, H., Ito, J., and Mester, A., editors, *Prosody and Prosodic Interfaces*, pages 63–94. Oxford University Press, Oxford.
Zsiga, E. and Zec, D. (2013). Contextual evidence for the representation of pitch accents in Standard Serbian. *Language and Speech*, 56(1):69–104.
Zuraw, K. and Hayes, B. (2017). Intersecting constraint families: an argument for Harmonic Grammar. *Language*, 93(3):497–548.

A. Prose sources

Title	Author	Year	Genre	N sentences	N decasyll. sentences	N poetry-like sentences
Glava šećera	Laza Lazarević	1875	short story	639	74	6
Večiti mladoženja	Jakov Ignjatović	1878	novel	3,344	104	39
Prva brazda	Laza Lazarević	1885	short story	179	10	4
Hajduk Stanko	Janko Veselinović	1896	novel	8,691	420	146
Nečista krv	Borisav Stanković	1910	novel	3,384	75	30
Seobe	Miloš Crnjanski	1929	novel	2,964	27	11
Na Drini ćuprija	Ivo Andrić[23]	1945	novel	5,283	74	31
Gospodjica	Ivo Andrić	1945	novel	3,031	56	27
Koreni	Dobrica Ćosić	1954	novel	6,452	260	95
Seobe 2	Miloš Crnjanski	1962	novel	17,877	394	161
Derviš i smrt	Meša Selimović	1966	novel	7,256	208	75
Rani jadi	Danilo Kiš	1970	short story collection	1,114	27	8
Omerpaša Latas	Ivo Andrić	1977	novel	4,271	63	18
Besnilo	Borislav Pekić	1983	novel	10,538	319	110
Hazarski rečnik	Milorad Pavić	1984	novel	4,126	52	13
Kandže	Marko Vidojković	2004	novel	4,266	73	27
			total	**83,415**	**2,236**	**801**

Table A.1.: Prose sources and by-source number of epic decasyllable-like sentences.

B. Ganging-up cumulativity and local constraint conjunction

Local constraint conjunction of two constraints assigns a violation to each candidate that violates both conjuncts within a designated local domain (in our case, the verse line). For example, HEAVY⇒STRONG & STRESS⇒(HIGH⇒STRONG) penalizes every verse line that simultaneously contains at least one misaligned heavy and at least one misaligned stressed high-toned syllable.

In strict-ranking OT, local conjunctions are employed to model all instances of ganging-up cumulativity. HG, including MaxEnt as its probabilistic implementation, captures most ganging-up interactions by default, thanks to the additivity of constraint violations. This has led many to argue that HG dispenses with the mechanism of constraint conjunction (Farris-Trimble 2008, Potts et al. 2010). However, HG still needs conjoined constraints to model a limited subset of ganging-up interactions, particularly those involving extreme superlinearity (Green and Davis 2014, Shih 2017, Milenković 2026b).

That being the case, we probed whether adding conjoined constraints improves our MaxEnt model's handling of ganging-up cumulativity in the BCMS folk epic. We analyze constraint pairs where independent violations of both conjuncts were possible, therefore excluding pairs in a superset relation, such as HEAVY⇒STRONG and STRESS⇒(HEAVY⇒STRONG). This leaves five constraint pairs listed in Table B.2. For each, we tested whether adding their conjunction improves model fit to the full dataset.

[23] Awarded the Nobel Prize in Literature in 1961.

We compared, using the LRT, the baseline MaxEnt model (five constraints) with five superset models, each augmented with one conjunction (six constraints). None of the conjoined constraints passed the LRT, as Table B.2 shows.

Conjunction	$\chi^2(1)$	p
Stress⇒(Heavy⇒Strong) & Stress⇒(High⇒Strong)	0.00002	0.996
Stress⇒(Heavy⇒Strong) & Stress⇒(Strong⇒Heavy)$_{\text{cadence}}$	0.00001	0.997
Stress⇒(High⇒Strong) & Stress⇒(Strong⇒Heavy)$_{\text{cadence}}$	1.02	0.31
Stress⇒(Strong⇒Heavy)$_{\text{cadence}}$ & Heavy⇒Strong	0.000005	0.998
Stress⇒(High⇒Strong) & Heavy⇒Strong	0.00002	0.997

Table B.2.: Likelihood ratio tests for conjoined constraints.