Studia Metrica et Poetica 5.2, 2018, 55–76

Testing the robustness of final strictness in verse lines

Varuṇ deCastro-Arrazola*

Abstract: In the field of metrics, it has long been observed that verse lines tend to be more regular or restricted towards the end (Arnold 1905). This has led to the Strict End Hypothesis [SEH], which proposes a general versification principle of universal scope (Hayes 1983). This paper argues that two main challenges hinder the substantiation of the SEH in a broad typological sample of unrelated verse corpora. First, the concept of strictness is too coarse and needs to be narrowed down to testable features or subcomponents. Second, explicit measures need to be developed which enable the systematic comparison of corpora, particularly when trying to capture potentially gradient features such as the relative faithfulness to a metrical template. This study showcases how to overcome these issues by analysing the entropy at different positions in the line for corpora in five languages (English, Dutch, Sanskrit, Estonian, Berber). Finally, I argue that, if the SEH is shown to be typologically robust, shared human cognitive features may provide a partial explanation for this puzzling asymmetry in verse lines.

Keywords: final strictness; verse universals; verse typology; cognition

Introduction

The words used in verse are subject to a number of constraints which are absent in everyday speech. By analysing how songs and poems are structured we can observe, for instance, that the discourse is organised into lines of similar length, that a pulse can be perceived by the regular alternation of strong and weak syllables, or that a number of adjacent lines end with exactly the same phonemes.

However, there seems to exist an asymmetry in the way these constraints are arranged: the beginning of lines are left relatively free, and later parts of the line are more constrained. This can be the result of constraints specific to the end of lines (e. g. rhyme), or due to some general constraint (e. g. alternate weak and strong syllables) being more stringent later in the line.

* Author's address: Varuṇ deCastro-Arrazola, Leiden University, Centre for Linguistics, Postbus 9515, 2300 RA Leiden, The Netherlands. E-mail: varunasarman@gmail.com.

https://doi.org/10.12697/smp.2018.5.2.03

The phenomenon has been most prominently mentioned with reference to early verse corpora, such as Ancient Greek (Prince 1989) or Sanskrit (Arnold 1905). However, it has also been noted that it is an "almost constant feature of numerous widely differing metrical systems of the world" (Kiparsky 1968: 138), which has led to the hypothesis that it "is in fact just a specific manifestation of a universal principle" (Hayes 1983: 388).

The Strict End Hypothesis (henceforth, SEH) covers a number of predictions, roughly summarised by the statement that "correspondence to a metrical pattern tends to be lax at the beginnings of units; strict at the ends" (Hayes 1983: 373). In order to verify the extent of the hypothesis, one needs to be able to verify whether a given sample of verse conforms to it or not. As I will argue in the following section (2), the concept of final strictness is too coarse to undertake a typological study. As defended by Bickel (2007: 247), the same problem applies to many traditional descriptive variables in linguistic typology (e. g. incorporation); instead of seeking universal definitions of such variables, it would be more productive, he maintains, to encode finer-grained variables: "such variables allow capturing rather than ignoring diversity, and they stand a greater chance to be codable in replicable ways across many languages".

Section 2 discusses the properties of a number of instances of final strictness from different languages, and proposes a formal division into two main classes of strictness. Beyond the characterisation of the sub-variables composing strictness in verse, there is also the issue of gradience: it is unfeasible to describe with precision a *relative* higher strictness in a *relatively* late position of a line without some sort of quantification. Section 3 showcases how a particular type of final strictness (the kind found in Vedic verse) can be quantified. This enables a finer-grained characterisation of the degree of strictness and finality, and allows systematic cross-linguistic comparisons. In the final discussion, I present a number of potential explanations for the different types of final strictness, and point to further ways of testing them empirically.

1. Types of strictness

The Strict End Hypothesis has a number of possible definitions, which means that it is flexible enough to cover a wide range of phenomena. However, if we want to systematically survey and compare like with like manifestations of the SEH, we need to be precise about the scope of the term. Characterising one or several types of final strictness becomes crucial if we want to (1) verify or falsify the universality of the phenomenon, and (2) investigate the possible causes of its pervasiveness. There are at least three problematic issues with the notion of SEH.

First, it is unclear whether the strictness applies only at one particular constituent level (e. g. the line), or to any kind of metrical constituent (e. g. stanzas, hemistichs, feet). This can be a source of confusion, since evidence and counter-evidence for the hypothesis can refer to completely different metrical domains. The article by Zwicky and Zwicky (1986) "Patterns first, exceptions later" offers a potential counter-example to the hypothesis, but in this case the strictness applies at the level of the stanza: limericks show greater regularity at the beginning lines. Still, most of the remarks on final strictness make reference to the level of the line. Besides, lines are possibly the only defining constituents of verse (Shapir 1995; Fabb 2015: 20), so it is a suitable domain for a working definition of SEH.

A second source of heterogeneity in the phenomena is where the focus of the strictness asymmetry is set: (1) an exceptional freedom at the very beginning, (2) an exceptional strictness at the very end, (3) a gradual increase of strictness (or decrease of freedom) encompassing the line as a whole. The trochaic inversion common in English iambic pentameter refers to an exceptional looseness at the beginning of the line (Hayes 1983). Rhyme constraints, on the other hand, usually target just the end of lines. Finally, other instances of SEH are reported to apply gradually, with strictness increasing from the beginning to the end of the line (e. g. Finnish *Kalevala* verse, Kiparsky 1968).

Third, in order to speak of strictness, one needs to posit a rule or restriction which can then be satisfied or violated; however, a variety of features (e. g. syllable weight, phonemes) can be subject to restrictions in verse templates. The cognitive representation of these features differ, and it is unclear whether strictness phenomena on different features can be directly compared as produced by a shared cause. Next, I review the main types of restrictions for which the SEH has been mentioned.[1]

[1] Several of the examples I mention are discussed by Fabb (2002), with additional analysis on the SEH.

2. Final strictness phenomena

2.1. Syllable type

The prototypical example of final strictness relies on the existence of a template which shows restrictions on syllable type. In Somali [Afro-Asiatic; Cushitic][2] *geeraar* verse lines, for instance, the third and fourth syllabic positions are required to be heavy and light respectively; the preceding positions, however, do not show a strong preference for either of the syllable types (Banti, Giannattasio 1996: 99). The same kind of stricter syllabic weight pattern is found in Sanskrit [Indo-European; Indo-Iranian] (Arnold 1905: 9). Languages where syllabic stress rather than weight is constrained in verse, e. g. English or Dutch [Indo-European; Germanic], also show lines with looser beginnings (de Groot 1936; Hayes 1983).

2.2. Syllable-to-position association

The number of syllables or morae associated to a metrical position is fixed in many traditions. However, certain deviations are permitted, such as the resolution, where a strong position, typically filled with a heavy syllable, is realised instead as two weak syllables. This kind of freedom is frequent in Greek [Indo-European; Graeco-Phrygian] iambic trimeter, but it does not occur in the final two strong positions (Prince 1989: 61). Similarly, in Somali *gabay* metre, the strong positions of the first half-line can be resolved as two light syllables; in the second half, however, only one of the three strong positions can be resolved, because the total number of syllables is fixed to six (Johnson 1996: 76). This means that the second half of the line is more restricted than the first. Another kind of Somali verse, *masafo*, constrains the number of morae per position in all but the very first position of the line, where 2, 3 or 4 morae can be realised (Banti, Giannattasio 1996: 94).

[2] In the first mention of a language in the current section, I include its genealogical information between brackets, extracted from Glottolog (Nordhoff, Hammarström 2011).

2.3. Word length

A number of verse traditions restrict the length of the line-final word, leaving the rest of words unconstrained in terms of length. In the Irish [Indo-European; Celtic] *Ae freislighe* quatrains, odd lines end in trisyllables, and even lines in dissyllables (Knott 1957: 13); Dyirbal [Pama-Nyungan] *Marrga* songs require line-final words to be bisyllabic (Dixon, Koch 1996: 181); in Finnish [Uralic; Finnic] *Kalevala* verse, monosyllables are "not permitted at the end of a line" (Kiparsky 1968: 138).

2.4. Phonemes

Rhyme is a very common feature in the poetry and singing of many languages. This involves limiting the choices of phonemes one can use at the end of lines; i. e. a relative free choice of phonemes in every position of the line, except at the end. Rhyme, hence, can be interpreted as a case of strict end.

Some languages, nonetheless, follow the opposite pattern: they restrict the phonemes occurring at the beginning of lines, and not elsewhere, by using line-initial alliteration. This type of strict beginning is attested throughout the Mongolic language family (Kara 2011, Krueger 1961), as well as in Welsh [Indo-European; Celtic] (Greenhill 2011). From a typological point of view, however, line-final rhyme is much more frequent than line-initial alliteration (Fabb 1999). This asymmetry further strengthens the SEH.

2.5. Words

Even more stringent are the verse templates where specific words or kind of words are required to close the line. In a Sardinian [Indo-European; Italic] *anninnia* from Bosa, each line in the song is closed by the formula *ninna ninna* (Sassu, Sole 1972: 121); in the Melpa [Nuclear Trans New Guinea; Central East New Guinea Highlands] *kang rom* style of songs, lines are composed by regular lexical words, but one of a restricted set of vocables (i. e. meaningless words) is produced at the end of each line (Niles 2011: 284). Further typological work is needed to assess whether these line-final formulae are more frequent than line-initial ones, as the latter are also attested in languages such as Kuna [Chibchan; Core Chibchan] (Sherzer 1982: 373).

2.6. Melody

The majority of the world's verse is performed in sung form; this usually entails the use of stable pitch classes (i. e. in the form of melodies). Robust typological evidence lacking, there are indications that fixed patterns of melodic pitch classes are more frequent at the end of lines than elsewhere. In the Nambudiri tradition of Veda recitation (in Sanskrit), lines are closed by a conventional melodic cadence over the last vowel or nasal consonant (Staal 1961: 50). Many verse systems show a similar pitch-cadence phenomenon, where a fixed (low) pitch closes each line (e. g. Huli [Nuclear Trans New Guinea; Enga-Kewa-Huli] songs, Pugh-Kitingan 1984: 107). While these traditions display particularly invariant melodic material in the *last* few syllables, others display an asymmetry by singing the *first* few notes with an undetermined, speech-like pitch, while the rest of the line employs stable notes (e. g. Tedaga [Saharan; Western Saharan] songs, Brandily 1976: 176).

2.7. Rhythm

A related line-final effect is the lengthening of the very last syllable of the line. This is observed e. g. in Warlpiri [Pama-Nyungan; Desert Nyungic], Somali or Sardinian (Banti, Giannattasio 1996; Sassu, Sole 1972; Turpin, Laughren 2013), and has probably a widespread typological distribution (Nettl 1956: 66). One of the results of lengthening is that the temporal interval between the last syllable of a line and the first of the following line is increased with respect to the preceding inter-syllabic intervals. This same effect is also achieved by a simple pause, or by leaving an empty beat, as seen at the end of lines in many nursery rhymes, such as *Eeny, meeny, miny, moe*, where each strong syllable is followed by a weak one except at the end of lines, or *Hickory dickory*, where all lines except the third show a line-final empty beat. This kind of truncation is used preferentially at the end of lines, as it provides a structural pause (Hayes, Kaun 1996; Hayes, MacEachern 1996).

2.8. Template and instance strictness

All these phenomena can be interpreted as evidence for the SEH in one way or another. However, I want to argue that it is useful to be specific about the kind of features which are constrained, and to characterise the stringency of the constraints.

Some of the cited examples consist of completely stringent, fixed phenomena which mark the right edge of lines somehow, e. g. by requiring the presence of specific phonemes, words or pitches. We can refer to them as examples of *template strictness*.

The constraints related to syllable type and syllable-to-position association, on the other hand, usually exhibit varying degrees of stringency. We can call them examples of *instance strictness*. Requiring a specific rhyme to close the line is arguably a categorical feature encoded in the verse template; the increasing consistency in using light or heavy syllables, though, is a gradual effect observed when a collection of instances of the same template are analysed.

This binary classification of final strictness phenomena can already be helpful in order to better understand their possible cause or function. Nevertheless, quantifying the stringency of any kind of strictness is still crucial if one wants to verify whether it is categorically localised at one of the edges, whether it is gradual, and, if gradual, the extent to which the different positions in the line are restricted. The next section develops a case study where the instance strictness of syllable type is quantified and compared across verse samples in five different languages. This is the prototypical case of final strictness, and, given its gradual nature, it constitutes a suitable object for quantitative examination.

3. Measuring strictness

3.1. Materials

We analyse data from five languages: three Indo-European languages from two different branches (English and Dutch from the Germanic branch, and Sanskrit from the Indo-Iranian branch), the Uralic language Estonian, and the Afro-Asiatic language Tashlhiyt Berber. To be sure, the sample is not broad enough to make strong claims about the universality of final strictness; however, I describe a methodology which can be easily extended to include further data in future studies. The choice of languages attempts to maximise the typological coverage, while being constrained by the availability of sizeable digitized corpora.

The kind of strictness being measured here is the one about restrictions on syllable type (cf. Section 2.1). Hence, any verse sample where at least certain positions require a particular syllabic feature (e. g. weight or stress) are suitable for the analysis. Table 1 lists the five languages used in the analysis, together with the total number of lines, and the number of samples (e. g. authors) for each language.

The Sanskrit sample includes lines from the *Ṛgveda*, a text composed in the second millennium BC, probably earlier than 1200 BC (Witzel 1995). The templates used are quantitative, i. e. they contrast heavy and light syllables. For the current analysis I have employed the summary statistics provided by Gunkel and Ryan (2011), who list the proportion of heavy syllables for each position in metres of eight, eleven and twelve syllables. All three metres follow a general iambic pattern. However, that the verse lines from the *Ṛgveda* show final strictness has been known for a long time: "in all metres the rhythm of the latter part of the verse is much more rigidly defined than that of the earlier part" (Arnold 1905: 9). The present analysis involves almost thirty-eight thousand lines of verse.

The Dutch sample includes 9079 lines by two 20th-century poets: J. P. Kal (b. 1946) and C. O. Jellema (1936–2003). Most of the lines belong to 14-line-long sonnets and follow an iambic pentameter template. For purposes of syllable-position identification, lines longer or shorter than ten syllables have been excluded from the original corpus. Each line has been automatically scanned using a scansion algorithm (van Oostendorp, 2014) which takes into account the syllable's lexical stress and its environment, and yields a binary result for each syllable:

Table 1: Summary of corpora used in the analyses

Language	**ISO**	**Lines**	**Samples**	**Family**	**Branch**
English	eng	4198	2	Indo-European	Germanic
Estonian	est	8811	20	Uralic	Finnic
Dutch	nld	9079	2	Indo-European	Germanic
Sanskrit	san	37908	3	Indo-European	Indo-Iranian
Tashlhiyt Berber	shi	314	7	Afro-Asiatic	Berber

metrically stressed (1) or unstressed (0). The iambic pentameter template predicts that odd positions will contain unstressed values, and even positions stressed values. The iambic pentameter template predicts that odd positions will contain unstressed values, and even positions stressed values.

The English sample contains 4198 lines by John Milton (1608–1674) and William Shakespeare (1564–1616). Our analysis is based on the digital text annotated by Bruce Hayes, which assigns one of four stress levels to each syllable: 0 = unstressed, 1 = secondary stress, 2 = primary, 3 = phrasal (Hayes, Wilson, Shisko 2012). Given that most accounts of English verse only distinguish a binary opposition of stress, I have collapsed values 1, 2 and 3 into a single category of stress, opposed to unstressed (0). The lines used for the analysis all follow an iambic pentameter template. The sample by Shakespeare is derived from his 154 sonnets, excluding sonnet 145 which is composed in iambic tetrameter, and excluding lines longer than 10 syllables (i. e. ending in feminine rhyme). The sample by Milton is drawn from books 9 and 10 of his work *Paradise Lost*, also applying the filter to retain only 10-syllable-long lines.

The Estonian sample summarises 8811 lines composed by 20 different authors from the late 19th and early 20th centuries. The data are taken from statistics provided by M.-K. Lotman and M. Lotman (2013), where each syllable is assigned a stress value ranging from 0 (unstressed) to 4 (phrasal stress). As with the English sample, only a binary distinction between stressed and unstressed has been retained. All lines follow a trochaic tetrameter template, where odd positions generally contain stressed syllables, and even positions unstressed.

The Tashlhiyt sample contains 314 lines of verse belonging to seven different songs. Each song follows a different template, but all of them are quantitative; hence, positions are expected to contain either a heavy or a light syllable (as in the Sanskrit corpus). The song texts, their scansion and thorough analyses have been published by Dell and Elmedlaoui (2008).

In all five corpora we observe templates where two classes of syllables are used in a controlled way. Still, the nature of these two classes depends on the phonological features of each language, and the interpretation of which syllables constitute deviations from the template depends on the method of analysis used for each corpus. In the Estonian corpus, the alternating syllable classes are based on word stress, which is always word-initial in this language (Harms 2017). The syllable classes in the English and Dutch corpora are also based on word stress, but this feature plays a more important role than in Estonian, since its placement within the word is not fully predictable (Van der Hulst 1984). On the other hand, the Sanskrit and Tashlhiyt syllable classes are not based on stress but on weight, where syllables ending in a coda (and/or in a long vowel for Sanskrit) are considered heavy (Arnold 1905; Dell, Elmedlaoui 2002).

In terms of the kinds of methods used to detect deviant syllables, the Dutch corpus differs from the other four in that for each syllable the scansion algorithm takes into account the neighbouring syllables and the ideal metrical template to determine its prominence value. In the other corpora, whether a syllable is considered weak or strong does not depend on the metrical context, but is determined exclusively on linguistic grounds.

3.2. Statistical analyses

For each syllabic position in the corpora as coded here, a binary feature (0, 1) indicates the prominence value (related to stress or weight) for that syllable. It is assumed all lines in a sample follow the same template, which regulates the placement of stress or weight. In order to measure the consistency of this regulation, I compute the entropy for each position of the template using the standard equation by Claude Shannon (1948):

(1)

$$entropy = -\sum_{i=0}^{n} p_i \log_2 p_i$$

For a given syllabic position in a template, indicates the proportion of syllables with feature, where can be either 1 (heavy or stressed), or 0 (light or unstressed). Given the binary prominence feature used here, the formula yields an entropy value between 0 and 1, with lower values indicating that the position is more consistent in using a particular prominence value. If e. g. the second position in a set of iambic pentameter lines always contains a stressed syllable, its entropy would be the lowest possible, i. e. 0. The highest uncertainty (entropy = 1), instead, would be obtained if half the lines had a stressed syllable, and half the lines an unstressed syllable, indicating effectively that that position in the line is unregulated.

This measure does not capture higher order dependencies between positions, such that an unstressed syllable in the second position of an iambic pentameter is likely preceded by a stressed syllable (i. e. trochaic inversion). Nevertheless, it does allow us to analyse the relative unigram consistency related to syllable type regulation across line positions, and across corpora.

As explained in the previous section (3.1), the different corpora here analysed contain lines of various lengths. These range from a minimum of eight syllables (e. g. in the Estonian tetrameter), to a maximum of fourteen syllables (in one of the Tashlhiyt songs). In order to assess the effect of syllable position on the position's entropy in a unified way, it becomes necessary to normalise the line lengths of the different templates. Hence, instead of using the original line positions, I have computed relative syllable positions (*relpos*) which fall within the range [0,1]. For a line with *n* number of syllables, the relative syllable position for position *x*is computed with the formula (x−1)/(n−1). As a result, the *relpos* for the first syllable of a tetrameter is (1−1)/(8−1)=0/7=0, while the eighth and last syllable would yield (8−1)/(8−1)=7/7=1.

In order to assess whether the relative syllabic position (*relpos*) predicts a decrease in entropy (i. e. less heterogeneity towards the end of the line), I fit a mixed-effects model to the whole set of samples combined. The different language corpora are added as random factors (see the complete model under Equation 2 as called in `R` using the `lme4` package, Bates, Mächler, Bolker, Walker 2015). The first term on the right part of the equation (after the) indicates the extent to which the entropy for a given metrical position is explained by its relative position within the line (*relpos*). The part in parentheses models the extent to which the effect of *relpos*, and the effect of a position being weak/strong (*prominence*) depend on the language of the corpus. For some languages, the *relpos* effect may be larger/smaller, and weak or strong positions may have a higher/lower overall entropy. In order to assess with more detail the robustness of the SEH for each individual language, I also perform separate linear regressions for each language corpus.

(2)

$$entropy \sim relpos + (1 + relpos + prominence|language)$$

3.3. Results

For reproducibility purposes, the data for the five corpora, as well as the `R` scripts used to perform the analyses, are included at the online Supplementary Information.[3]

Figure 1 plots the entropy values for each syllabic position within each sample, grouped by language. Separate regression lines are drawn for positions treated as strong and weak in their respective templates.

[3] https://github.com/vdca/seh1.

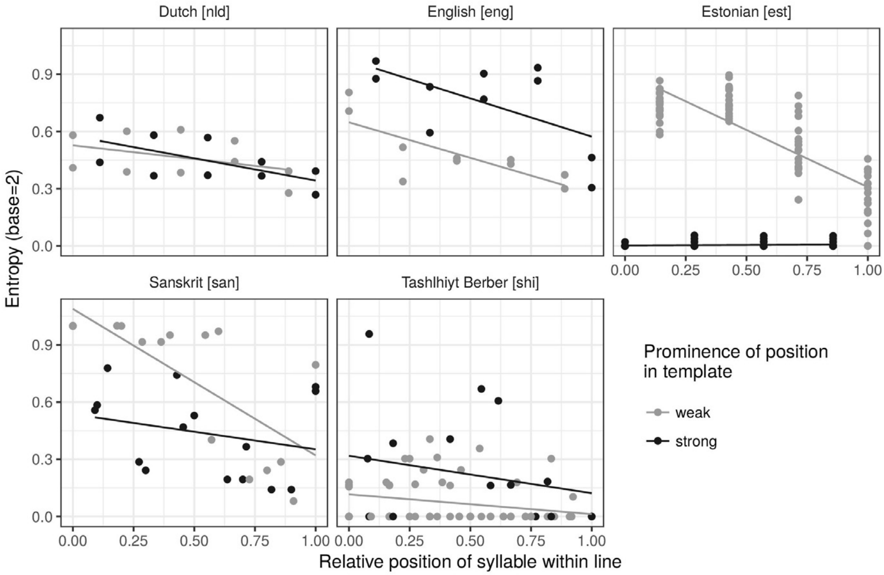

Figure 1. Entropy of syllabic feature values (stressed *vs.* unstressed; heavy *vs.* light) for each position in each sample of verse

We can observe a general downtrend in entropy for all languages, suggesting that the initial hypothesis that there is an increase in consistency in the use of weight and stress is correct. In the Estonian sample, the trend for strong positions is completely flat, with no apparent increasing strictness. In fact, deviations from the ideal template (stressed syllables on odd positions) are close to null. The Sanskrit sample also shows a flatter and lower regression line for strong positions. However, this is not the case for the other three languages.

Visual inspection reveals that the final position of Sanskrit verse breaks with the increasing strictness tendency, a phenomenon well described in the literature (Arnold 1905: 112). Furthermore, a relative increase in entropy towards the middle of the line (e. g. English and Sanskrit) suggests the presence of division of the line in two half-lines.

Table 2 shows the results of the mixed effects model fitted to the data. Relative position within the line (*relpos*) proves to be a good predictor of entropy, with a strong negative estimate indicating a decrease in uncertainty as the syllable position increases. In the first row of the table (Intercept), the Estimate indicates what the entropy would be for the initial position of a random corpus of verse, regardless of its language. In the second row (*relpos*), the Estimate indicates how much that entropy value would decrease on average as one goes from the line-initial position to the line-final position of the template. A comparison of this model with the corresponding null model without the fixed predictor (i. e. *relpos*) indicates that the prediction of entropy significantly improves by adding relative syllable position as a predictor (χ^2= 9.63, Pr (>χ^2) = 0.0019).

Figure 2 shows how these estimates are to be adjusted for each of the language corpora; that is, it visualizes the random part of the mixed model under Equation 2. The vertical dashed line represents the average, language-neutral results, i. e. those shown under Table 2). Red/blue figures indicate that a language shows a lower/higher value with respect to the language-neutral baseline.

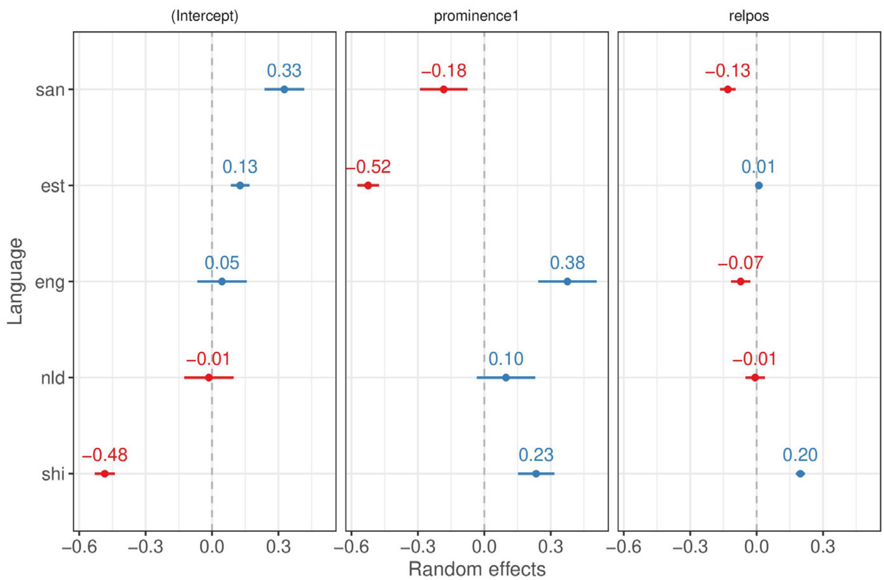

Figure 2. Random effects for the five different languages in the sample. The plot indicates the amount by which the model estimates are adjusted according to the language of the data under analysis. Negative adjustments with respect to the baseline are plotted in red.

In the leftmost panel (Intercept) we observe that Sanskrit shows the highest overall entropy, and Tashlhiyt Berber the lowest, confirming the plots in Figure 1. In the second row of Table 2 we had observed that later syllables have a negative estimate, i. e. they are predicted to have a lower entropy. The rightmost panel (*relpos*) reveals that Sanskrit is the language where this decreasing-looseness effect is strongest, i. e. it shows the steepest slope from high entropy at the beginning of the line, to low entropy at the end. Finally, the central panel (prominence1) displays the varying effect of prominence on entropy. Here, we corroborate that strong positions in Estonian (and in Sanskrit to a lower extent) are much stricter than weak positions (i. e. are expected to show a lower entropy), whereas the opposite effect is found in English.

Individual linear regressions per language sample (Table 3 of the Supplementary Information) confirm that the robust overall decrease in entropy also holds for each language corpus. The prominence value of the position (whether weak or strong) further improves the model for all languages except Dutch, where weak and strong positions follow a similar downtrend.

Table 2. Results of the full mixed model, with relative syllable position as the fixed predictor, and random slope and intercept for the effect of prominence and syllable position in each language

	Term	**Estimate**	**Std. Error**	**t value**	**Pr (> t)**
1	(Intercept)	0.575	0.117	4.91	0.00804
2	relpos	-0.305	0.0616	-4.95	0.00995

Discussion

Binary restrictions on the type of syllable used at certain positions of the line are common to many verse traditions. As proposed by e. g. Jakobson (1966), metrical rules are often regarded as binary; nevertheless, Ryan (2011) argues that this binarity appears to be gradient. Moreover, the direction of the gradience can be predicted to a certain extent using the relative position of the syllable. In the samples of verse here analysed, there is a robust trend for the binary restriction on weight or stress to be more stringent later in the line. This supports a specific type of the SEH, namely, if a verse tradition regulates a syllabic feature, it will be more consistent the later the syllable occurs in the line.

Visual inspection of the Estonian corpus suggested a second predictor of entropy: prominence. Even though four out of five corpora show distinct entropy profiles for weak and strong positions, there is no overall prediction: in some cases strong positions show higher entropy (English, Tashlhiyt), and in others, lower entropy (Estonian, Sanskrit). A note of caution is needed when using position prominence as a predictor of entropy. At least two potential confounds can drive the direction of the effect of prominence on entropy: (1) the structure of the lexicon, (2) the coding algorithm.

In a language with higher proportion of light syllables, consistently using light syllables in weak positions proves easier, hence strong positions are predicted to have a higher entropy (i. e. it is more likely that they contain light syllables, than the other way round). On the other hand, as explained in the description of the materials, the algorithms used to decide the weight/stress feature of a syllable differ, e. g. in Estonian, even the lightest degree of stress is coded as stressed; in Dutch, the neighbouring syllables are considered in order to maximise the fit to the ideal template. The facts expressed in those two caveats hinder a direct interpretation on the predictive effect of position prominence on entropy. Further research is needed which takes lexico-statistical data from the feature of interest as a predictor of the effect of prominence.

If a robust typological tendency is established, we can hypothesise that it derives from features shared by the whole population under study, like some aspect of basic human cognition. In Section 2 I propose that the diverse phenomena under the general SEH can be subdivided into two main types: template strictness, and instance strictness. The contrast between the categorical nature of the first type, and the gradual nature of the second suggests different causes. From a general perspective, I propose that categorical asymmetries encoded in the templates (e. g. rhyme, cadence) may have a facilitating *function*, and that the gradual asymmetries (e. g. selection of syllable type) are a *result* of some cognitive bias.

A plausible, low-level cognitive bias is the gradual increase of attention as new temporal stimuli are processed. Each verse template displays a regular alternation of features, but the regularity gets a stronger representation as the line develops and more syllables satisfy the template, as proposed by the Bayesian predictive coding framework (Vuust, Witek 2014). Alternatively, if the creator or recipient of the verse lines entrains to some regular temporal sequence (e. g. of syllables or beats), the dynamic attending theory (Jones, Moynihan, MacKenzie, Puente 2002) predicts that clusters of neurons will synchronise with that regularity. The synchronisation strengthens as more stimuli are processed, and, if attention peaks correlate with neuronal firing, one is expected to be extra sensitive by the end of the line.[4]

Hypotheses on increase of attention can be suitable for gradual phenomena, and particularly for loose beginnings, where full-entrainment has not yet taken place. Nevertheless, it does not fit well with categorically final phenomena, such as rhyme. A straightforward interpretation of these is that they work as

[4] These arguments could also be applied to larger chunks of verse, such as couplets or stanzas. However, there is good evidence that lines are treated as whole units in working memory (Fabb 2014), making it a suitable candidate for the unit of attention increase.

boundary-markers, making the constituent structure of verse easier to parse. However, an equivalent explanation is available for left-boundary markers. Hence, a number of functional accounts specific to template final strictness can be put forward, but they all remain tentative in the lack of robust empirical work, probably experimental.

A feature common to all the categorical asymmetries I have discussed in Section 2 is that they limit the number of choices at the very end of the line, by restricting the choice of phonemes, the word-length, or requiring a specific closing-word. This effectively reduces the cognitive load of the performer. Still, this cognitive advantage at the end of lines can be seen as serving various alternative *functions*.

The first is proposed, for instance, by Niles (2011) when discussing the fixed vocables closing lines of *kang rom* Melpa songs: "because of their regularity in delimiting a line of text, these vocables perhaps also allow the performer a brief rest and chance to mentally compose his or her thoughts for the next line". This explanation may be particularly relevant for improvised traditions, where the following line needs to be composed while still singing the current one; yet, it also applies to non-improvised performances, where saving cognitive load can facilitate the recall of the next line (Rubin 1995). Similar functions have been attributed to the pervasive final lengthening observed in everyday speech (Fletcher 2010).

A related advantage of marking right boundaries with predictable material is that it enables a smoother turn-taking. The previous hypothesis worked best in the context of solo performance; however, there are traditions where two poets engage in dialogue-like exchange of lines (Egaña 2007). In this context, one needs to compose a line *while* the other is singing. Hence, predicting the end of a line gives the poet some advantage in order to plan the next line and execute it without delay. Again, this has a parallel in everyday speech, where it has been shown that the gap between turns is so short, that speakers must plan in advance and accurately predict the end of the interlocutor's utterance (Stivers et al. 2009).

An alternative proposal by Fabb (2014) is that reduced cognitive load at the end of a line facilitates the recall of earlier elements within that same line. The hypothesis relies on the idea that lines are processed as single units within working memory. The final part of the line already has a recency-effect advantage, and having a reduced cognitive demand at the end would leave more room for keeping in mind earlier linguistic content. Earlier in this section, I have argued that final strictness may ease the planning of the *following* line, while Fabb proposes that final strictness eases the remembering of the *current* line.

Given the diversity of final strictness phenomena, it is unlikely that they all have the same cause or function, so apparently contradictory hypotheses may in fact prove complementary. In this specific case, Fabb's proposal covers better the function of final strictness during verse perception, while the planning hypothesis applies to the process of verse production, as exemplified by the quote on Melpa singing (Niles 2011). In order to test the coverage of these hypotheses, behavioural experiments can be conducted which manipulate line-final elements and compare recall and reaction times. Moreover, the predominant ecological context where verse is created and consumed in a given tradition may also require dissimilar cognitive explanations; in particular, the visual aid available when composing texts in written form, such as the Dutch or Estonian samples here analysed, would need to be taken into account.

Finally, the pervasive presence of predictable material at the end of lines can have aesthetic reasons. As argued by Huron (2006), a source of pleasure in music lies in the fulfillment of expectations. Moreover, added aesthetic value may be produced when expectations are violated or kept on hold for some time, and satisfying them. Although gathering empirical evidence of aesthetic value poses methodological challenges, physiological measures such as skin conductance response (Mas-Herrero, Zatorre, Rodríguez-Fornells, Marco-Pallarés 2014) can potentially be used in order to test whether line-final fixed elements are a more significant source of aesthetic pleasure compared to predictable material on other locations in the line.

Conclusion

The Strict End Hypothesis has been discussed in the field of metrics for more than a century. Furthermore, it has been posited that it may be universal in nature. In this paper I address two challenges which precede the typological verification of the claim. First, I argue that a range of diverse phenomena are categorised within the SEH, making the hypothesis effectively intractable. At the very least, two types of strictness should be distinguished: template and instance strictness. Second, in order to measure the extent and degree of each type of strictness, quantification or other kinds of fine-grained description are necessary. I showcase this by characterising gradual strictness in a dataset including verse from five languages, using entropy as a proxy for strictness. Finally, I discuss a number of testable cognitive explanations of the phenomena, which move towards the goal of understanding the *why* of final strictness[5].

[5] The paper has benefited from insightful comments by Teresa Proto, Marc van Oostendorp and an anonymous reviewer. The usual disclaimers apply. This research was made possible thanks to the project Knowledge and culture (Horizon grant 317-70-010) funded by NWO (Dutch Organisation for Scientific Research).

References

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

Banti, Giorgio; Giannattasio, Francesco 1996. Music and metre in Somali poetry. In: Hayward, Richard J.; Lewis, Ioan M. (eds.), *Voice and power: Essays in honour of B. W. Andrzejewski* (*African Languages and Cultures*, Supplement 3). London: School of Oriental and African Studies, 83–127.

Bates, Douglas; Mächler, Martin; Bolker, Ben; Walker, Steve 2015. Fitting linear mixed-effects models using lme4. In: *Journal of Statistical Software* 67(1), 1–48.

Bickel, Balthasar 2007. Typology in the 21st century: Major current developments. In: *Linguistic Typology* 11(1), 239–251.

Brandily, Monique 1976. Un chant du Tibesti (Tchad). In: *Journal des Africanistes* 46(1), 127–192.

de Groot, Albert Willem 1936. De structuur van het vers. In: *De Nieuwe Taalgids* 30, 197–212.

Dell, François; Elmedlaoui, Mohamed 2002. *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer Academic Publishers.

Dell, François; Elmedlaoui, Mohamed 2008. *Poetic meter and musical form in Tashlhiyt Berber songs*. Köln: Rüdiger Köppe.

Dixon, Robert M. W.; Koch, Grace 1996. *Dyirbal song poetry. The oral literature of an Australian rainforest people*. St. Lucia: University of Queensland Press.

Egaña, Andoni 2007. The process of creating improvised bertsos. In: *Oral Tradition* 22(2), 117–142.

Fabb, Nigel 1999. Verse constituency and the locality of alliteration. In: *Lingua* 108, 223–245.

Fabb, Nigel 2002. *Language and literary structure: The linguistic analysis of form in verse and narrative*. Cambridge: Cambridge University Press.

Fabb, Nigel 2014. The verse-line as a whole unit in working memory, ease of processing, and the aesthetic effects of form. In: Currie, Greg; Kieran, Matthew; Meskin; Aaron; Moore, Margaret (eds.), *Royal Institute of Philosophy Supplement 75*. Cambridge: Cambridge University Press, 29–50.

Fabb, Nigel 2015. *What is poetry? Language and memory in the poems of the world*. Cambridge: Cambridge University Press.

Fletcher, Janet 2010. The prosody of speech: Timing and rhythm. In: Hardcastle, William J.; Laver, John; Gibbon, Fiona E. (eds.), *The handbook of phonetic sciences*. West Sussex, UK: Blackwell Publishers, 523–602.

Greenhill, Peter 2011. Bardic rhythm: The implications from *cerdd dant* studies. In: *Studia Celtica* 45, 131–153.

Gunkel, Dieter; Ryan, Kevin 2011. Hiatus avoidance and metrification in the *Rigveda*. In: Jamison, Stephanie W.; Melchert, H. Craig; Vine, Brent (eds.), *Proceedings of the 22nd annual UCLA Indo-European conference*. Bremen: Hempen, 53–68.

Harms, Robert T. 2017. *Estonian grammar*. London: Routledge.

Hayes, Bruce 1983. A grid-based theory of English meter. In: *Linguistic Inquiry* 14(3), 357–394.

Hayes, Bruce; Kaun, Abigail 1996. The role of phonological phrasing in sung and chanted verse. In: *The Linguistic Review* 13, 243–303.

Hayes, Bruce; MacEachern, Margaret 1996. Are there lines in folk poetry? In: Hsu, Chai-Shune K. (ed.), *UCLA Working Papers in Phonology* 1, 125–142.

Hayes, Bruce; Wilson, Colin; Shisko, Anne 2012. Maxent grammars for the metrics of Shakespeare and Milton. In: *Language* 88(4), 691–731.

Huron, David 2006. *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: Massachusetts Institute of Technology Press.

Jakobson, Roman 1966. *Selected writings*. Vol. IV: *Slavic epic studies*. The Hague, Paris: Mouton.

Johnson, John William 1996. Musico-moro-syllabic relationships in the scansion of Somali oral poetry. In: In: Hayward, Richard J.; Lewis, Ioan M. (eds.), *Voice and power: Essays in honour of B. W. Andrzejewski* (*African Languages and Cultures*, Supplement 3). London: School of Oriental and African Studies, 73–82.

Jones, Mari Riess; Moynihan, Heather, MacKenzie, Noah; Puente, Jennifer 2002. Temporal aspects of stimulus-driven attending in dynamic arrays. In: *Psychological Science* 13(4), 313–319.

Kara, György 2011. Alliteration in Mongol poetry. In: Roper, Jonathan (ed.), *Alliteration in culture.* London: Palgrave Macmillan, 156–179.

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

Knott, Eleanor 1957. *An introduction to Irish syllabic poetry of the period 1200–1600*. Dublin Institute for Advanced Studies.

Krueger, John Richard 1961. *Poetical passages in the Erdeni-yin Tobči: a Mongolian chronicle of the year 1662 by Sayang Sečen*. ’s-Gravenhage: Mouton.

Lotman, Maria-Kristiina; Lotman, Mihhail 2013. The quantitative structure of Estonian syllabic-accentual trochaic tetrameter. In: *Trames* 17(3), 243–272.

Mas-Herrero, Ernest; Zatorre, Robert J.; Rodríguez-Fornells, Antoni; Marco-Pallarés, Josep 2014. Dissociation between musical and monetary reward responses in specific musical anhedonia. In: *Current Biology* 24, 1–6.

Nettl, Bruno 1956. *Music in primitive culture*. Cambridge: Harvard University Press.

Niles, Don 2011. Metric melodies and the performance of sung tales in the Hagen area. In: Rumsey, Alan; Niles, Don (eds.), *Sung tales from the Papua New Guinea highlands. Studies in form, meaning, and sociocultural context*. Canberra: Australian National University Press, 275–201.

Nordhoff, Sebastian; Hammarström, Harald 2011. Glottolog/Langdoc: Defining dialects, languages, and language families as collections of resources. In: Kauppinen, Tomi; Pouchard, Line C.; Keßler, Carsten (eds.), *Proceedings of the first international workshop on linked science 2011*. Vol. 783. http://ceur-ws.org/Vol-783/paper7.pdf

Prince, Alan 1989. Metrical forms. In: Kiparsky, Paul; Youmans, Gilbert (eds.), *Rhythm and meter*. San Diego, CA: Academic Press, 45–80.

Pugh-Kitingan, Jacqueline 1984. Speech-tone realisation in Huli music. In Kassler, Jamie C.; Stubington, Jill (eds.), *Problems & solutions: Occasional essays in musicology presented to Alice M. Moyle*. Sydney: Hale & Iremonger, 95–120.

Rubin, David C. 1995. *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford: Oxford University Press.

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

Sassu, Pietro Constantino Alberto; Sole, Leonardo 1972. Funzione degli stereotipi nel canto popolare sardo. In: *Rivista Italiana di Musicologia* 7(1), 115–144.

Shannon, Claude E. 1948. A mathematical theory of communication. In: *Bell System Technical Journal* 27(3), 379–423.

Shapir, Maksim 1995. “Versus” vs “prosa”: prostranstvo-vremja poeticheskogo teksta. In: *Philologica* 2(3/4), 7–47.

Sherzer, Joel 1982. Poetic structuring of Kuna discourse: The line. In: *Language in Society* 11, 371–390.

Staal, Johan Frederik 1961. *Nambudiri Veda recitation*. Hague: Mouton.

Stivers, Tanya; Enfield, Nick J.; Brown, Penelope; Englert, Christina; Hayashi, M.; Heinemann, Trine; Hoymann, Gertie; Rossano, Federico; de Ruiter, Jan; Yoon, Kyung-Eun; Levinson, Stephen C. 2009. Universals and cultural variation in turn-taking in conversation. In: *Proceedings of the National Academy of Sciences of the United States of America* 106(26), 10587–10592.

Turpin, Myfany; Laughren, Mary 2013. Edge effects in Warlpiri *yawulyu* songs: Resyllabification, epenthesis, final vowel modification. In: *Australian Journal of Linguistics* 33(4), 399–425.

Van der Hulst, Harry 1984. *Syllable structure and stress in Dutch*. Dordrecht: Foris.

van Oostendorp, Marc 2014. Introducing a scansion machine for Dutch poetry and prose. In: *Loquens* 1(1), 1–7.

Vuust, Peter; Witek, Maria A. G. 2014. Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. In: *Frontiers in Psychology* 5, Article 1111.
https://www.frontiersin.org/articles/10.3389/fpsyg.2014.01111/full

Witzel, Michael 1995. Early Sanskritization. Origins and development of the Kuru state. In: *Electronic Journal of Vedic Studies (EJVS)* 1(4), 1–26.

Zwicky, Arnold M.; Zwicky, Ann D. 1986. Patterns first, exceptions later. In: Channon, Robert; Shockey, Linda (eds.), *In honor of Ilse Lehiste*. Dordrecht: Foris, 525–537.

Addendum. Supplementary Information

Table 3. Results of the linear model applied to each language corpus, with position prominence and relative syllable position as predictors.

	Language	Term	Estimate	Std. Error	*t* value	Pr (> *t*)
1	eng	(Intercept)	0.654	0.0668	9.79	2.1e-08
2	eng	relpos	-0.386	0.106	-3.64	0.00204
3	eng	prominence1	0.312	0.0678	4.6	0.000258
4	est	(Intercept)	0.735	0.0248	29.6	3.24e-66
5	est	relpos	-0.297	0.034	-8.73	3.5e-15
6	est	prominence1	-0.603	0.0223	-27.1	5.31e-61
7	nld	(Intercept)	0.547	0.0464	11.8	1.29e-09
8	nld	relpos	-0.189	0.0737	-2.56	0.0204
9	nld	prominence1	0.00408	0.0471	0.0866	0.932
10	san	(Intercept)	0.966	0.0915	10.6	2.91e-11
11	san	relpos	-0.504	0.143	-3.53	0.00147
12	san	prominence1	-0.257	0.091	-2.83	0.00859
13	shi	(Intercept)	0.127	0.034	3.73	0.000351
14	shi	relpos	-0.125	0.0552	-2.26	0.0264
15	shi	prominence1	0.158	0.0431	3.65	0.000459