

Short Article

Cite this article: Nagy, Ben, Artjoms Šeļa, Mirella De Sisto and Petr Plecháč. 2025. "Metronome: tracing variation in poetic meters via local sequence alignment." *Computational Humanities Research*, 1:e1, <https://doi.org/10.1017/chr.2025.1>

Received: 16 October 2024
Revised: 24 January 2025
Accepted: 26 February 2025

Keywords:
computational poetics; distant reading; meters; sequence alignment; versification studies

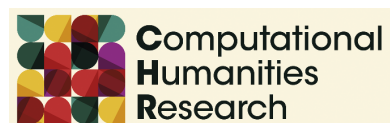
Corresponding author: Mirella De Sisto;
Email: M.DeSisto@tilburguniversity.edu

B.N. and A.Š. are joint first authors.

*This article has been updated since original publication. A notice detailing the change has also been published.

  This Short article was awarded Open Data and Open Materials badges for transparent practices. See the Data availability statement for details.

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



Metronome: tracing variation in poetic meters via local sequence alignment*

Ben Nagy¹, Artjoms Šeļa^{1,2} , Mirella De Sisto³  and Petr Plecháč²

¹Institute of Polish Language, Polish Academy of Sciences (IJP PAN), Krakow, Poland; ²Institute of Czech Literature, Czech Academy of Sciences, Prague, Czechia and ³Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, The Netherlands

Abstract

All poetic forms come from somewhere. Prosodic templates can be copied for generations, altered by individuals, imported from foreign traditions, or fundamentally changed under the pressures of language evolution. Yet, these relationships are notoriously difficult to trace across languages and times. This article introduces an unsupervised method for detecting structural similarities in poems using local sequence alignment. The method relies on encoding poetic texts as strings of prosodic features using a four-letter alphabet; these sequences are then aligned to derive a distance measure based on weighted symbol (mis)matches. Local alignment allows poems to be clustered according to emergent properties of their underlying prosodic patterns. We evaluate method performance on a meter recognition tasks against strong baselines and show its potential for cross-lingual and historical research using three short case studies: (1) mutations in quantitative meter in classical Latin, (2) European diffusion of the Renaissance hendecasyllable and (3) comparative alignment of modern accentual-syllabic meters in 18–19th century Czech, German and Russian. We release an implementation of the algorithm as a Python package with an open license.

Plain language summary

Metrical poetry (poetry with a designated pattern of stresses, lengths, or some other arranging principle: as opposed to “free verse”) is connected to both culture and language. Different kinds of meters carry individual meanings and associations, layered on top of the words – certain meters might suggest “heroic” or “funny” or “impressive.” Poems, and so meters, spread among people (or, more broadly, between cultures), and meters change over time; we wanted to study this. However, this is difficult because of the huge differences in how meters work across history and between languages. Since these ideas of change and spread suggest an evolutionary metaphor, we tried using techniques from biology and represented meters like DNA sequences. This turned out to work well. In this article, we explain how and why, as well as presenting some interesting findings drawn from very different kinds of metrical poetry throughout history. We show that our method is sensitive to minute metrical differences by tracing an odd poem by the ancient poet Catullus. It also proves useful in detecting broad structural changes in a meter that spread across a dozen European languages during the Renaissance. Finally, we show it can simultaneously handle different modern meters in different languages, paving the way to a global history of poetic forms.

Introduction

All poetic forms come from somewhere, even if, perhaps because of their familiarity, it is natural to think they have been with us forever. Indeed, a prosodic template can be copied in a culture for centuries or millennia; but it can also be altered by individuals, imported from foreign traditions, adapted from vernacular folk songs, or fundamentally changed by the pressures of language evolution. Yet these relationships are notoriously difficult to trace across languages and times: a lot of evidence is too fragmentary; it depends too much on written heritage, or modern ethnographic observations. National schools of metrical studies are also disparate, and rarely agree with each other, while literary scholarship over the course of the 20th century distanced itself from metrics, which are now more closely associated with linguistics and phonology. Research on the history of versification that was built on top of historical linguistics in the works of Antoine Meillet (1923), Roman Jakobson (1929) and, later, Mikhail Gasparov (1996), was in decline, but interest is presently resurgent, following the expansion of computational and quantitative methods that allow new angles of inquiry (De Sisto 2023; Polilova 2018; Šeļa, Plecháč, and Lassche 2022).

This work is a continuation of the historical line of thinking about meters and metrical variation, and aims primarily to describe relationships between the *forms* of poetic texts. We

use local sequence alignment to identify regions of structural similarities in poems; the method relies on encoding poetic texts as strings of prosodic features using a simple four-letter alphabet that can be used with any language and is independent of the organizing verse principle (quantitative, syllabic, accentual). These sequences are then aligned to derive a distance measure based on weighted symbol (mis)matches. The resulting relationships do not inherently imply any evolutionary connection between texts, but the structural similarity is illuminating: it might identify and localize emerging metrical derivatives, show continuity in forms across times and traditions, or signal formal descent, which can be useful in historical studies. While the prosodic and structural features of verse today are increasingly used for authorship attribution (Nagy 2021; Plecháč 2021), this work focuses on detecting patterns that transcend individual variation and shape cultural history on a large scale: the often unseen, but powerful force of poetic meters.

Poetic meters organize speech into periodic units, causing utterances to be perceived in systemic relation to each other (Frog 2021). Meters can be defined based on various prosodic systems, the most familiar being patterns of stressed and unstressed syllables (accentual-syllabic), but they can also be defined based on syllable length (quantitative meters, as in classical Latin), or simple syllable count (isosyllabic meters in Romance languages). In addition, the prosody is only one component of what some might call a “meter.” Old English verse specifies both a prosodic pattern and also a system of alliteration, or (better known) many rhyming forms are defined by both prosody and sonic correspondences.

Meters act as rules, with varying levels of fuzziness and variation, undoubtedly constrained by cognition and languages’ natural prosody (deCastro-Arrazola 2018a; 2018b; Rubin 1995). These rules, even when scholars and poets formulate them prescriptively, depend on shifting social conventions and are always tested against the affordances of a language (Shapir 2000). This entails variation in the poetic realization of abstract patterns: extra-metrical stresses, additional syllables and other irregularities test the boundaries of metrical inertia and audience expectation. Metrical patterns are reaffirmed by individual usages that copy forms from the past: taken together, poems reveal governing tendencies that are not easily overwritten by individual flair. This important cultural aspect of metrical forms has prompted research into bottom-up approaches to meter recognition and verse regularity (Plecháč and Birnbaum 2023; Porter 2018; Šeļa and Gronas 2022) that supports clustering according to the shared structural principles that emerge from data. This descriptive approach might be considered a counterpoint to generative frameworks that abstract metrical variation under the premise of “deep structures” and language-specific “correspondence rules” (Fabb and Halle 2008), drawing away from the historical and cultural component of poetic meters.

We evaluate our method on the non-trivial task of meter recognition, using a labelled cross-linguistic corpus, showing the method’s ability to recognize similarity in metrical organization simultaneously across different languages. Classification, however, is not our primary goal: we highlight the method’s potential for historical research in three case studies that span multiple languages, versification systems and time periods (from classical Latin quantitative verse to accentual-syllabic Romantic poetry). In the hope that the methodology will find further use in the community, we release an accompanying *metronome* library for Python.¹

Method

The guiding metaphor for the development of the *Metronome* tools was the idea of the genome, and the four DNA molecules that are the basis for all cellular life. While consisting of only four symbols, genomes vary hugely in length and are capable of expressing minor variation, such as among human siblings, or vast differences – the same symbols define a cactus, a dolphin or a hummingbird. In the same way, the alphabet for the metronome was chosen to be as simple as possible, allowing the same kinds of analyses that are performed with DNA; comparison between very disparate samples, grouping related families into clades or clusters, and expressing or locating changes over time. We use the following symbols: for strong and weak syllables, “S” and “w,” for word breaks, “.” and for end of line “|” (the specific glyphs are only mnemonics, any would work). For example, the metronome of a line “Once upon a midnight dreary, while I pondered, weak and weary” might be expressed as S . wS . w . Sw . Sw . S . w . Sw . S . w . Sw . |. Defining prosody in this way (strong/weak vs stressed/unstressed) allows the same alphabet to encode quantitative traditions, like classical Latin, as well as modern accentual traditions. Note that, for simplicity, we mostly rely on orthographic boundaries between words, not on phonological units. However, this is not a requirement for the method to work; it depends on the encoding choices: for traditions where elision is important, such as classical Latin or some Romance forms, a word break can be omitted in the metronome representation – see, e.g., Note 3 for how this was handled for Renaissance meters.

In bioinformatics, there are several strategies for sequence alignment, depending on the task. For genomes (or metronomes) that are very different in length and possibly very dissimilar, the best way to compare sequences is with a “local sequence alignment.” This approach finds the shared region between two sequences that is the most similar. In contrast, a “global sequence alignment” is a better choice where the strings are closely related. We apply the well-known Smith–Waterman algorithm (Smith and Waterman 1981) for local sequence alignment, which uses a flexible alphabet and allows different bonuses and penalties for symbol (mis)matches. Smith–Waterman is agnostic regarding the sequence alphabet – it is commonly applied both to nucleic acid sequences (four symbols) as well as proteins (20 symbols), and produces a match score that depends on a bonus/penalty matrix. In our work, the best local match score is normalized using the self-score of the shorter sequence, producing a comparable distance (or similarity) in [0,1].

The general intuition behind the bonuses and penalties (which determined the initial values) was based on empirical understanding of poetic meter – for example, matching line lengths are rewarded more than matching syllable strengths, and mismatches between word breaks and syllables are not penalized at all (although matching word breaks are rewarded, which allows the algorithm to group schemes that have customary caesura positions, discussed more in Appendix A). To optimize the bonus/penalty matrix, we then performed a series of iterative classification tasks on the cross-language dataset, and visually inspected the clustering results of several corpus subsets with known metrical schemes. The metronome package leverages methods from the BioPython package (Cock et al. 2009) which provides highly optimized routines for the sequence alignments. Our code adds an assortment of sampling strategies, with the core of the package being a parallel method that produces an $n \times n$ distance matrix from n input samples (which have been encoded as metronomes).

¹ Available at <https://github.com/bnagy/metronome-paper>.

These distance matrices can be used directly as inputs to create hierarchical clustering / dendrogram methods (such as `hclust` in R) or to spatial clustering methods like UMAP (using a “pre-calculated” metric).

Evaluation

We evaluate the performance of the algorithm in a non-trivial cross-linguistic task by looking at clustering performance on samples of individual poems in twelve meters, and coming from four languages (Czech, German, Russian and Classical Latin). We compare this performance to three baselines: simple edit distance, local alignment with an unweighted substitution matrix, and a machine-learning approach based on the frequencies of prosodic n -grams.

Data and sampling

Our modern language subset (67,940 works) covers the six most common meters: iambic and trochaic pentameter, tetrameter, and tetrameter/trimeter (alternating lines of tetrameter and trimeter, sometimes called “common meter” in English). The Czech subset (44,416 works), German subset (15,172 works) and Russian subset (8,352 works) come from the PoeTree dataset (Plecháč et al. 2023). Metrical annotation in Czech was performed using machine learning methods, and the output was manually verified (Plecháč 2016); in German it was produced using a rule based algorithm (Bobenhausen and Hammerich 2015), and a similar approach was implemented in Russian, and used the accentual classes of words to infer rhythm. The Classical Latin subset (2,588 works) covers the six most common meters from that corpus: elegiac couplets, hendecasyllables and stichic hexameter (making up the bulk of samples and authors), as well as a few senarii, scazons and alcaics. The largest language / meter set is Czech iambic pentameter (over 17,000 works) and the smallest is Latin alcaics (38 works, all from Horace).

Since the samples are of such disparate size, we downsampled to just 10 examples of each language / meter combination, yielding 240 random samples per evaluation run. We then classified samples by meter (twelve meter labels) which implies a random baseline accuracy of about 8 percent.

Classification methods

For the alignment-based methods, which operate in a similar way to Metronome, we created an $n \times n$ distance matrix (alignment distance), to which we applied a k -Nearest Neighbors (kNN) classifier ($k=7$). These alignment-based methods are as follows. For the Naïve sequence aligner, we take the best local alignment, but without permitting indels (insertions or deletions inside the aligned areas). This is similar to a local Hamming distance, and is a very unforgiving alignment (since word breaks are included), used only as a baseline. Next we added indels to the Naïve aligner, making it more like a local Levenshtein distance but with uniform match/mismatch bonuses and penalties – this is similar to the way local sequence aligners would work on DNA sequences. Finally, we used the full Metronome configuration with a variable penalty matrix. The fourth method, and the only one that was not alignment-based, was a standard machine-learning approach, in which we applied an SVM classifier to the frequencies of the 500 most-frequent symbol 9-grams, using an 80/20 train/test split. Nine-grams were chosen because the symbol alphabet is very small, and so it is advantageous to the classifier to be able to match

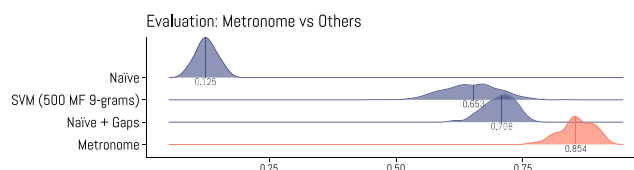


Figure 1. Classification performance of Metronome vs SVM and two baseline alignment algorithms.

Note: Smoothed distribution of accuracy results over 50 random subsamples, with median scores.

almost an entire line. Based on testing, the best results were with 500 most-frequent nine-grams, reduced by SVD to 50 dimensions and then z-scaled. It must be noted here that the SVM classifier is hampered greatly by the presence of word-breaks. For most meters, word breaks are not part of the metrical constraints, they are only relevant to some questions of style and in a few traditions that have mandatory or customary caesurae in certain places (it is very important in classical Latin and also in modern alexandrines, for example). In other words, in terms of predicting the name of the metrical pattern, the word breaks are noise. If the aim were simply to create an optimal classifier, machine-learning approaches perform at least as well as Metronome (and are much faster), at the expense of nuance and interpretability, by simply discarding the word breaks.

Results

The classification tests were run 50 times each, and the full distributions and medians are reported in Figure 1. As discussed above, the unexpectedly poor performance of the SVM classifier (usually a solid benchmark) is mainly due to the presence of word breaks, but it is also clear that the custom penalty matrix used in Metronome provides a solid increase in performance across languages and traditions, compared to the standard string aligner. Classification accuracy, however, is merely a proxy for the real utility of Metronome as a tool to understand metrical *relationships* between poems. In the following section we provide a few small examples of this potential.

Showcases

Showcase A: Mutations in classical Latin meter

The first showcase simply highlights the sensitivity of the metronome to minor metrical variations. During a manual inspection of the dendrogram of the poems of Catullus, it was observed that one hendecasyllable poem, Catullus *Carmen* 55, was quite different to all of the others of that meter. This can be seen in Figure 2 – notice how poem 55 branches off from the rest of the hendecasyllable group much “earlier” (indicating more difference), and yet is still clearly part of the overall clade. Thinking that this might be an algorithmic error, we inspected the text, and discovered that 55 is a metrical oddity. In 55, Catullus often collapses the central double breve of the choriamb (— ∪ —) to form a molossus (— — —), yielding a line of ten syllables instead of the standard 11 (hendecasyllable is, literally, “11 syllables”). This variation can be seen in Figure 3, which shows the start of 55, alongside its metronome transformation, contrasted with 41, a normal hendecasyllable with a standard choriamb in every line. This license occurs only twice more in all of Catullus (in 58b),

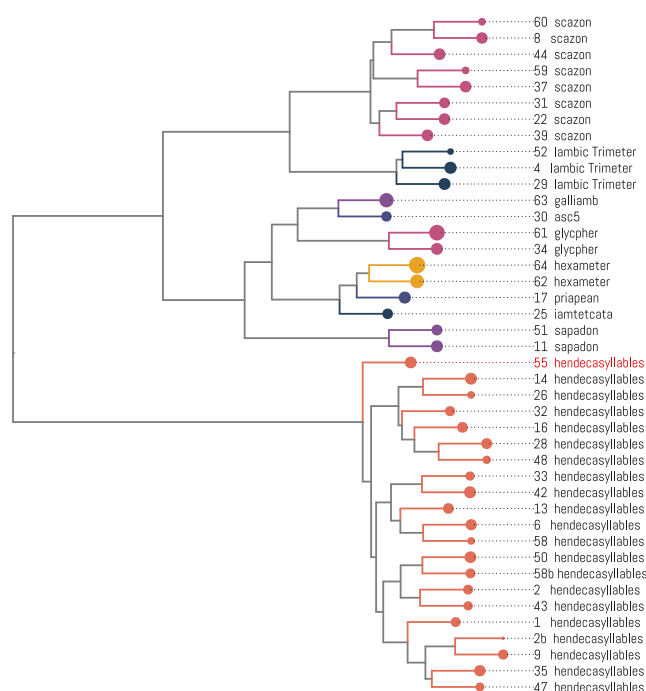


Figure 2. Cladogram of a selection of poems by Catullus. *Carmen* 55, while still composed in hendecasyllables, is visibly different to the rest of that clade.

Oramus, si forte non molestum est,	SSS.S.Sw.S.wS.S.
demonstres ubi sint tuae tenebrae.	SSS.wS.S.wS.wS.S.
Te Campo quaeisivimus minore,	S.SS.SSwS.wS.S.
te in Circo, te in omnibus libellis,	.S.SS.w.wS.wS.wS.S.
te in templo summi Iovis sacrato.	S.SS.SS.wS.wS.S.
In Magni simul ambulatione	S.SS.wS.wS.wS.S.
femellas omnes, amice, prendi,	SSS.SS.wS.wS.S.
quas vultu vidi tamen sereno.	S.SS.SS.wS.wS.S.
...	
Ameana puella defututa	SwSw.wSw.SwSS.
tota milia me decem poposcit,	Sw.Sww.S.wS.wS.S.
ista turpiculo puella naso,	Sw.SwwS.wS.wS.S.
decoctoris amica Formiani.	SSSw.wSw.SwSS.
Propinqui, quibus est puella curae,	wS.wS.wS.wS.wS.S.
amicos medicosque convocate:	wS.wS.wSw.SwSS.
non est sana puella, nec rogare	S.S.Sw.wSw.S.wS.S.
qualis sit solet aes imaginisum.	SS.S.wS.wS.wS.S.

Figure 3. A visual comparison of the metronome strings (formatted to add line breaks) for the beginning of *Carmina* 55 (variant with collapsed choriamb) and 41 (standard hendecasyllable).

making 55 a very odd animal indeed.² It is precisely this kind of “loose” family grouping, based on the meters as they are written, not as they are defined by textbooks, that highlights the exciting potential of the metronome as a tool to provide insights from a “distant reading” of poetic meter.

The raw data for this showcase were provided by David Chamberlain, who maintains an extensive digital selection of fully scanned Greek and Latin verse under an open license (Chamberlain 2023). Minor post-processing was all that was required to convert it to the Metronome string format.

² An approachable description of the meters of Catullus, including a discussion of this issue, can be found in Peter Green’s bilingual edition (Green 2005) p. 32 ff. It has been noted that poem 58b does not cluster apart from the rest of the hendecasyllables; although this has not been carefully investigated, it is presumed that the two instances of the metrical irregularity are not enough of a ‘mutation’ to be noticeable to the algorithm.

Showcase B: European diffusion of a “Renaissance meter”

The second showcase focuses on a typological comparison of different implementations of Renaissance meter. During the Renaissance, a new metrical form spread across Europe by replacing previous forms, or by causing readjustments to the pre-existing templates. The meter was characterized by an obligatory stress on the 10th syllable of a line; in Italian, because of the penultimate stress in words, lines usually ended with one additional unstressed syllable, making an 11-syllable line. This meter – the endecasillabo – became widely known in Europe through Petrarch’s poetry. Italian, and likely French, isosyllabic examples later shaped the staple meter of modern English: the iambic pentameter.

In each tradition that adopted the innovation, the meter underwent some changes and adaptations to the recipient language and versification style which led to different implementations of the same poetic form (De Sisto 2020). The new meter probably originated in the Occitan tradition (Beltrami 1986; Billy 2000; Di Girolamo and Fratta 1999) and first reached Italy from there. The Europe-wide renown of Petrarch and his *endecasillabo* quickly spread the new meter to other Romance traditions, like Spanish, and vernaculars of the Italian peninsula, like Neapolitan, Venetian and Sicilian. French poetry was already using decasyllabic verse (*amour courtois*) (Hudson 1919), but this form was adjusted to the Italian trend of composing sonnets in *endecasillabo* (Hudson 1919; Key, Cave, and Bowie 2006), before being replaced by the alexandrine meter. Occitan verse also influenced Catalan and Portuguese poetry, which already had a pre-existing decasyllabic form.

Later in the Renaissance, this new metrical form spread independently to English and to Dutch poetry. Subsequently, the Dutch Renaissance meter influenced German and Frisian poets. All West Germanic adaptations developed strict iambic cadence from the contact with isosyllabic forms. This was the beginning of modernity in European poetry: the rise of accentual-syllabic meters everywhere east of France.

The metrical alignment (see Figure 4) of small samples from many Renaissance traditions clearly distinguishes between the three previously described ‘directions’ of adoption of the form: (1) Italian peninsula and Spanish; (2) Occitan, Catalan and French; and, finally, (3) Germanic traditions that pivoted to regular iambic meter (the values on the graph show how ‘free’ the samples are with regard to the placement of stresses – note the low values in the Germanic samples, suggesting the prominence of the accentual pattern). While the finer historical relationships between traditions may remain opaque to this method (which depends on the individual samples, and technical issues with word boundaries like synalepha),³ overall it captures the key structural and linguistic similarities emerging from the implementation of the same form.

The data employed in this showcase were originally compiled by Mirella De Sisto (2020); the samples were selected from the works of poets who can be considered representative of the Renaissance tradition of their respective languages. Manual annotation was

³ Poetry in many Romance traditions (Spanish, Italian ...) makes extensive use of synalepha – merging two syllables into one metrical syllable – which complicates syllable counting and word boundary placement. We manually annotated all instances of synalepha, and treated cross-word syllable elision as one ‘phonetic word’, placing a word boundary after the contracted sequence. For metronome analysis, all synalephas were annotated as a single syllable: “strong” if there was a strong syllable present on either side of the elision, and “weak” when both syllables were unstressed.

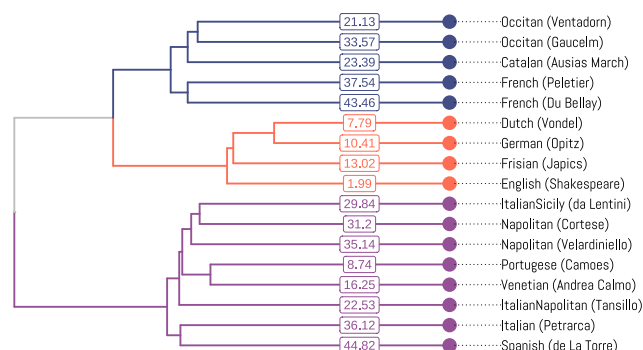


Figure 4. A metronome-based cladogram of various samples of Renaissance meter. Note: The inset number is the entropy-based variability from the regular metrical form (see Šeĵa and Gronas 2022). Shakespeare is the most regular, de La Torre the least.

performed, in collaboration with experts in the phonology and metrics of the respective languages.

Showcase C: Modern expansion: accentual-syllabic verse

During the 18–19th centuries, foot-based accentual-syllabic verse, which emerged through interaction with isosyllabic Renaissance meter, rapidly spread from Germany to the North and East (Gasparov 1996, 207–9; Kazartsev 2022) and was critical for establishing newly emerging national literatures. Through highly influential (and highly edited) folk song collections (Abrahams 1993; Leerssen 2012), like the *Volkslieder* (1786), regular trochaic meters became closely associated with oral traditions throughout Europe, while iambic verse, from German courtly odes to Shakespeare’s dramatic genius, provided a basis for high-status secular literature. This formed a cornerstone for the enduring cultural – and, as a result, semantic – opposition of iambic and trochaic meters (Šeĵa, Plecháč, and Lassche 2022).

As with any technology, conquest and power played a major role in metrical adoption, and it is not a coincidence that “fashionable” accentual-syllabic verse spread quickly through the territories, communities and languages of the Austro-Hungarian and Russian empires, also following the unification of literary languages and school systems (Gasparov 1996, 238–9). New verse often disrupted already existing versification, or weakened local alternatives (for example, syllabic verse in Polish, Czech and Ukrainian).

The last showcase (Figure 5) demonstrates the possibility of cross-linguistic alignment of distinct modern poetic forms in Czech, German and Russian. We are able to recognize the structural unity of the same meters (colors), detect the broader divide between trochee and iamb (inner and outer regions on the plot), yet preserve language-specific variation (point shapes) – the latter occurs partly because of differences in rhythm, and partly because word boundaries also encode differences in word lengths. German verse was a source for both Slavic literatures at different times, and the evidence here of structural similarity is a step toward better understanding the individual metrical lineages, and tracing formal connections through global literary history.

The data presented in Figure 5 are a subset of the same data we used in our evaluation (cross-linguistic meter recognition) and comes from annotated corpora of Czech, German and Russian poetry (all parts of the PoeTree).

Discussion

We move now to consider some limitations, and opportunities for further work. Our approach was mainly informed by poetic forms with isosyllabic meters and recurrent prosodic patterns, so we expect it to struggle with forms that emerged around other principles. For example, alliteration patterns that are part of the versification system in Old English and Old Norse (*dróttkvtt*) cannot currently be traced and would require a different approach to encoding the “alphabet” for alignment. Tonic verse, which can be composed of lines of different syllable lengths, regulated by a constant number of stressed positions, might also be difficult to analyze, since the algorithm is sensitive to line length. The same can be said of poetic forms in the same meter that are distinguished based on the shape of stanza and/or a pattern of rhymes. For example, both *ottava rima* and English sonnets can be written in iambic pentameter, yet represent distinct traditions. In this case, the expansion of the alphabet would be straightforward: adding a symbol for a stanza boundary would allow us to trace basic stanza composition (but make the analysis less broadly applicable). Additionally, the general methodology of Metronome (alignment of abstracted schemes) might be applied to rhyme patterns instead of rhythm sequences: one can envision, for example, a reconstruction of the history and typology of European sonnets based on classic rhyme-to-letter encoding: *abab|abab|cdcd|cdcd*.

Our current approach is also under-investigated in relation to caesurae – fixed-position word boundaries that can play a significant role in defining meters, like the French alexandrine (a meter of 12 syllables, where a caesura consistently bisects the verse, 6+6). To address this, we performed metronome analysis on simulated variations of the alexandrine (see Appendix B), to see if caesura-based forms can be distinguished from plain syllabics in our current approach. Preliminary results are very promising: even under conservative restrictions, the metronome tends to distinguish pseudo-poems that are governed by caesurae vs pseudo-poems of the same line-length that are not.

Finally, the method depends on the availability of raw scansions of large collections of poetry. This can be tricky, since scansion itself depends on existing metrical theory, shared within a community, a consensus that is rarely maintained across national traditions of scholarship. On one hand, theory-agnostic scansion is hardly ever possible, but on the other hand, the presence of theory in automated scansion systems can create self-fulfilling cycles, where empirical patterns are forced to fit a few anticipated schemes (for review see De Sisto et al. 2024), which regularizes and conceals actual variation.

In the authors’ other work we often advocate for the utility of bottom-up approaches to meter recognition, since any pattern that is repeated enough times, any strong organizational principle, will always make itself visible; the very nature of systematic meter will cause it to leave its mark on any observation or measurement, be it taken on a text’s prosody, morphology, or syntax (Gasparov and Tarlinskaja 2008). Meters are, essentially, quite simple technologies of rule iteration: the more you repeat one, the more obvious it becomes, even for such crude measuring instruments. This makes even a “sub-optimal” automated scansion, made without expert knowledge of the underlying theory and tradition, usable and useful.

In conclusion, this study shows how the alignment of metrical strings can be enlightening in a variety of contexts – from tracing slight deviations in well-established meters, to recognizing large

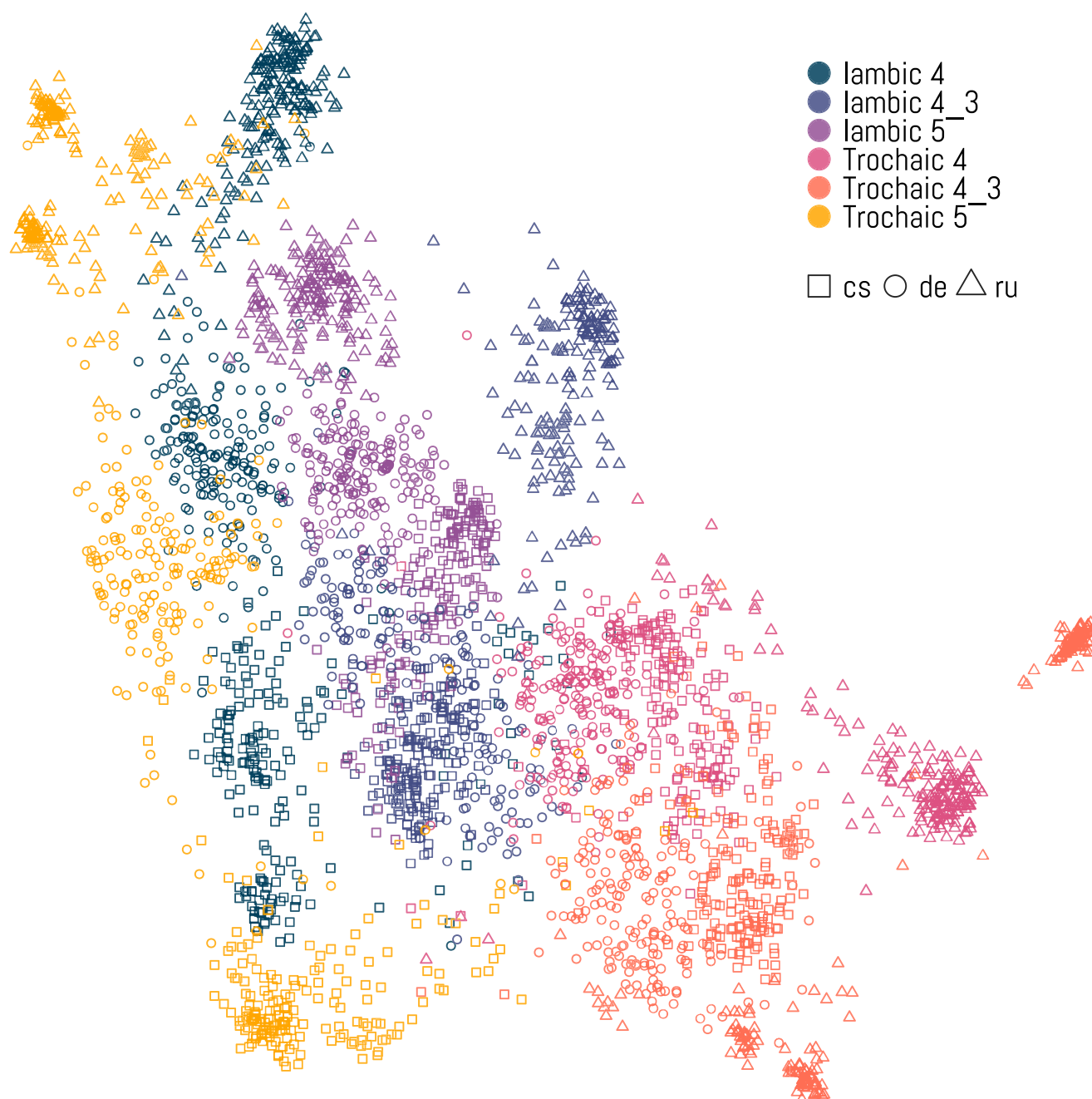


Figure 5. UMAP cluster of 3222 poems in Czech, German and Russian from the PoeTree corpus, in the six most common European meters.
 Note: Metronome distance is used as the clustering metric.

groups of similar forms across languages and times. Our proposed method works well with both long and short sequences, and is applicable in a wide range of poetic traditions (our showcases feature three major ones: quantitative, syllabic and accentual-syllabic). The method's sensitivity to large and small structural differences in verse organization highlights its potential in both historical and cross-linguistic comparative research.

Data availability statement. The repository may be found at <https://github.com/bnagy/metronome-paper>. All code and data are available under CC-BY, except where restricted by upstream licenses. The code repository includes full reproduction data and code for the evaluation, as well as various supplemental figures and explanations.

Acknowledgements. Data visualisations are produced with *ggplot* (Wickham2016) and cladograms use *ggtree* (Yu et al. 2016).

Author contribution. B.N. wrote the metronome software package and performed the evaluation and Classical Latin case-study. A.Š. conceptualized use cases for metrical alignment, provided a theoretical framework and was responsible for the accentual-syllabic showcase. P.P. prepared the modern language corpus. M.D.S. was responsible for the 'Renaissance' meter showcase. All authors contributed equally to the preparation and editing of the final article.

Funding statement. B.N. was supported by Poland's National Science Centre, project 2020/39/O/HS2/02931. A.Š. was supported by the project "Large-Scale Text Analysis and Methodological Foundations of Computational Stylistics"

(SONATA-BIS 2017/26/E/HS2/01019). M.D.S. was supported by the research project CIGE/2022/114 funded by the Valencian Government (GE 2023, Conselleria d'Innovació, Universitats, Ciència i Societat Digital). P.P. was supported by the project "European Poetry: Distant Reading (2023–2025)" (Czech Science Foundation GA23-07727S).

Competing interests. The authors declare none.

Ethical standard. The authors affirm this research did not involve human participants. No artificial intelligence (AI) tool was employed in this research nor during the writing process.

References

- Abrahams, Roger D.** 1993. "Phantoms of Romantic Nationalism in Folkloristics. Publisher: University of Illinois Press." *The Journal of American Folklore* 106, no. 419: 3–37. <https://doi.org/10.2307/541344>.
- Beltrami, Pietro G.** 1986. "Cesura Epica, Lirica, Italiana: Riflessioni Sull'endecasillabo di Dante." *Metrica* 4: 67–107.
- Billy, Dominique.** 2000. "L'invention de l'endecasillabo." In *Carmina Semper et Citharae Cordi. Études de Philologie et de Métrique Offertes à A. Menichetti*, edited by Gérard-Zai, Marie-Claire et al., 31–46. Geneva: Slatkine.
- Bobenhausen, Klemens, and Benjamin Hammerich.** 2015. "Métrique Littéraire, Métrique Linguistique et Métrique Algorithmique de l'Allemand Mises en Jeu Dans le Programme Métricalizer2." *Langages* 199, no. 3: 67–88. <https://doi.org/10.3917/lang.199.0067>.
- Chamberlain, David.** 2023. "Hypotactic: Greek and Latin meter." Accessed April 17, 2024. <https://hypotactic.com/>.
- Cock, Peter J. A., Tiago Antao, Jeffrey T. Chang, Brad A. Chapman, Cymon J. Cox, Andrew Dalke, Iddo Friedberg, Thomas Hamelryck, Frank Kauff, Bartek Wilczynski, and Michiel J. L. de Hoon.** 2009. "BioPython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics." *Bioinformatics* 25, no. 11: 1422–3. <https://doi.org/10.1093/bioinformatics/btp163>.
- De Sisto, Mirella.** 2020. "The Interaction Between Phonology and Metre. Approaches to Romance and West Germanic Renaissance Metre." PhD thesis, Radboud University.
- De Sisto, Mirella.** 2023. "The Development of a Poetic Tradition. A Study of a Dutch Renaissance Poetry Corpus." *Studia Metrica et Poetica* 10, no. 1: 36–68. <https://doi.org/10.12697/smp.2023.10.1.02>.
- De Sisto, Mirella, Laura Hernández-Lorenzo, Javier De la Rosa, Salvador Ros, and Elena González-Blanco.** 2024. "Understanding Poetry Using Natural Language Processing Tools: A Survey." *Digital Scholarship in the Humanities* 39, no. 2: 500–521. <https://doi.org/10.1093/llc/fqae001>.
- deCastro-Arrazola, Varun.** 2018a. "Testing the Robustness of Final Strictness in Verse Lines." *Studia Metrica et Poetica* 5, no. 2: 55–76. <https://doi.org/10.12697/smp.2018.5.2.03>. <http://ojs.utlib.ee/index.php/smp/article/view/smp.2018.5.2.03>.
- deCastro-Arrazola, Varun.** 2018b. "Typological Tendencies in Verse and their Cognitive Grounding." PhD diss., LOT. Accessed October 23, 2022. <https://www.lotpublications.nl/typological-tendencies-in-verse-and-their-cognitive-grounding>.
- Di Girolamo, Costanzo, and Aniello Fratta.** 1999. "I Decenari Con Rima Interna e la Metrica dei Siciliani." In *Dai Siciliani ai Siculo-toscani. Lingua, Metro e Stile, per la Definizione del Canone. Atti del Convegno di Lecce (21-23 Aprile 1998)*, edited by Rosario Coluccia and Riccardo Gualdo, 167–186. Galatina: Congedo.
- Fabb, Nigel, and Morris Halle.** 2008. *Meter in Poetry: A New Theory*. 1st edition. Cambridge: Cambridge University Press.
- Frog.** 2021. "Metrical Entanglement: The Interface of Language and Meter." In *Versification*, edited by Frog, Satu Grünthal, Kati Kallio, and Jarkko Niemi, 249–93. *Studia Fennica Litteraria*. Helsinki: Suomalaisen Kirjallisuuden Seura. <https://doi.org/10.21435/sflit.12>.
- Gasparov, M. L., and Marina Tarlinskaja.** 2008. "The Linguistics of Verse." *The Slavic and East European Journal* 52, no. 2: 198–207. <https://doi.org/10.2307/20459662>.
- Gasparov, Mikhail.** 1996. *A History of European Versification*. Oxford: Oxford University Press.
- Green, Peter.** 2005. *The Poems of Catullus: A Bilingual Edition*. Berkeley, Los Angeles and London: University of California Press.
- Hudson, William Henry.** 1919. *A Short History of French Literature*. London: G. Bell / Sons LTD.
- Jakobson, Roman.** 1929. "Zur Vergleichende Forschung über die Slavischen Zehnsilber." *Slavistische Studien* 1, no. 5: 7–20.
- Kazartsev, Evgeny.** 2022. "Iambic Verse in Different Literary Traditions. Publisher: De Gruyter (A)." *Glottology* 13, no. 1: 3–19. <https://doi.org/10.1515/glot-2021-2004>.
- Key, Sarah, Terence Cave, and Malcolm Bowie.** 2006. *A Short History of French Literature*. New York: Oxford University Press.
- Leerssen, Joep.** 2012. "Oral Epic: The Nation Finds a Voice." *Folklore and Nationalism in Europe During the Long Nineteenth Century* 4: 11–26.
- Meillet, Antoine.** 1923. *Les Origines Indo-Européennes des Mètres Grecs*. Paris: Les Presses universitaires de France. Accessed July 26, 2020. <http://archive.org/details/lesoriginesindoe00meiluoft>.
- Nagy, Benjamin.** 2021. "Metre as a Stylometric Feature in Latin Hexameter Poetry." *Digital Scholarship in the Humanities* 36, no. 4: 999–1012. <https://doi.org/10.1093/llc/fqaa043>.
- Plecháč, Petr.** 2016. "Czech Verse Processing System KVĚTA—Phonetic and Metrical Components." *Glottology* 7, no. 2: 159–174. <https://doi.org/10.1515/glot-2016-0013>.
- Plecháč, Petr.** 2021. *Versification and Authorship Attribution*. Prague: Charles University in Prague, Karolinum Press.
- Plecháč, Petr, and David J. Birnbaum.** 2023. "Rhythm-Based Authorship Recognition in Syllabic and Accentual-Syllabic Verse." *Literatura: Teoría, Historia, Crítica* 25, no. 2. <https://doi.org/10.15446/lthc.v25n2.108502>.
- Plecháč, Petr, Robert Kolár, Silvie Cinková, Artjoms Šeļa, Mirella De Sisto, Lara Nugues, Thomas Haider, and Neža Kocnik.** 2023. "PoeTree. Poetry Treebanks in Czech, English, French, German, Hungarian, Italian, Portuguese, Russian and Spanish [in ces]." Accessed April 16, 2024. <https://zenodo.org/records/10008459>.
- Polilova, Vera.** 2018. "Spanish Romancero in Russian and the Semantization of Verse Form." *Studia Metrica et Poetica* 5, no. 2: 77–108. <https://doi.org/10.12697/smp.2018.5.2.04>.
- Porter, J.D.** 2018. "The Space of Poetic Meter—Stanford Literary Lab." Accessed August 30, 2021. <https://litlab.stanford.edu/hooodistance/>.
- Rubin, David C.** 1995. *Memory in Oral Traditions: The Cognitive Psychology of Epic, Ballads, and Counting-Out Rhymes*. New York: Oxford University Press.
- Šeļa, Artjoms, and Mikhail Gronas.** 2022. "Measuring Rhythm Regularity in Verse: Entropy of Inter-Stress Intervals." In *CHR 2022: Computational Humanities Research Conference*, 231–42. Antwerp: CEUR-WS. https://ceur-ws.org/Vol-3290/short_paper5417.pdf.
- Šeļa, Artjoms, Petr Plecháč, and Alie Lassche.** 2022. "Semantics of European Poetry is Shaped by Conservative Forces: The Relationship Between Poetic Meter and Meaning in Accentual-Syllabic Verse. Publisher: Public Library of Science," *PLOS ONE* 17, no. 4: e0266556. <https://doi.org/10.1371/journal.pone.0266556>.
- Shapir, Maxim.** 2000. "Metrum et Rhythmus Sub Specie Semioticae." In *Univium versus*, vol. 1, 91–128. Moscow: Iazyki Russkoi Kultury.
- Smith, T. F., and M. S. Waterman.** 1981. "Identification of Common Molecular Subsequences." *Journal of Molecular Biology* 147, no. 1: 195–7. [https://doi.org/10.1016/0022-2836\(81\)90087-5](https://doi.org/10.1016/0022-2836(81)90087-5).
- Wickham, Hadley.** 2016. *Ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.
- Yu, Guangchuang, David Smith, Huachen Zhu, Yi Guan, and Tommy Lam.** 2016. "Ggtree: An R Package for Visualization and Annotation of Phylogenetic Trees with their Covariates and Other Associated Data." *Methods in Ecology and Evolution* 8: 28–36. <https://doi.org/10.1111/2041-210X.12628>.

Appendix A

Here, we briefly detail the custom bonus/penalty matrix that differentiates Metronome from the default Smith–Waterman alignment (in most descriptions of the algorithm this is referred to as a *substitution matrix*). For DNA sequences with the four labels A, G,

Table A1. The bonus/penalty matrix for metronome symbol (mis)matches

	—	w	S	.
—	3.5			
w	−4	2.5		
S	−4	−0.25	3	
.	−4	0	0	1

Note: The bonuses and penalties are symmetric, so only the lower triangular form is shown.

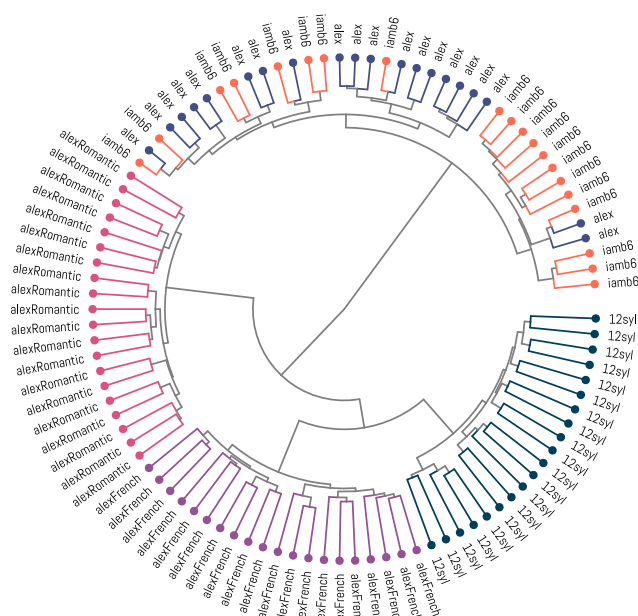
C, T, the default is to award 1 for a match and −1 for a mismatch. For protein sequence alignment, since some proteins are “closer” to others, there are range of different matrices to score pairwise matches from the 20 standard amino acids, but they need not concern us here. For our four metronome symbols, we began with the intuition that a (mis)match at the end of a line was the most important, since the algorithm will begin to favour pairs (of poems) whose lines have matching regular lengths. We then awarded a greater bonus for matches in the metrically strong syllable (S) than the weak (w), and eventually decided to award a small bonus (but no penalty) for matching word breaks (.) to help match poems with fixed or customary caesurae (Table A1). From there we performed a series of benchmarks, tuning the algorithm based on what looked to us like improvements in the way poems were grouped. This is a subjective process, using our own expertise and based on the corpora we have available – it is by no means unlikely that some analyses would benefit from tuning the default substitution matrix, and this can be easily done in the *metronome* software package.

Appendix B

Our analysis of modern meters in this article did not require the detection of caesurae – systemic word boundaries in the line (and often a syntactic pause) that help to define the meter in some traditions, like syllabic French verse and classical Latin. The standard modern example is the French 12-syllable alexandrine with a central caesura following a strong syllable. It is important to know whether the metronome can distinguish forms based on the presence, or placement, of caesurae. To test this, we generated pseudo-scansion sequences under four different conditions that imitate, albeit naïvely, verse prosody.

Each generated poem has a size drawn from a Poisson distribution with $\lambda = 14$, so that the length of all poems in the synthetic tests will average 14 lines, the length of a sonnet.

1. *Iambic hexameters with alexandrine caesura (alex)*: each line is strictly a sequence of “wSwSwS . wSwSwS,” with an obligatory word boundary after the 6th syllable. All other word boundaries in the line are determined randomly from a distribution of probable word-lengths;
2. *Plain iambic hexameter (Iamb-6)*: same as above, but no hard-coded caesura, word boundaries are assigned randomly;
3. *Classic French alexandrine (alexFrench)*: each line has a general structure of “xxxxxxS . xxxxxxS,” where x is any (w or S) symbol, and word boundaries are determined freely in each hemistich.

**Figure B1.** Clustering of simulated pseudo-poems that represent five conditions of the alexandrine form: three syllabic and four accentual-syllabic.

4. “Romantic” alexandrine (*alexRomantic*): has a general structure of xxxS . xxxS; this form is associated with Victor Hugo and Romantic experiments with the classic meter;
5. *Plain 12-syllable meter (12syl)*: a hypothetical meter that does not have any restrictions, except line length.

We generate lines under simple assumptions and do not follow the structure of any language in particular, nor do we try to plausibly reproduce word-length distribution and stress placement within words (so, e.g., unlikely words with prosodic structures like “ww” or “SSS” can occur). To distribute word boundaries within a line, we draw 1-, 2-, 3-, and 4-syllable words with the corresponding probability ratios 1:3:1:0.25. This ratio leads to a probability of 45 percent that there will be a word boundary after the 6th syllable in a non-alexandrine line, simply by chance. This is reasonably high, so the metronome needs to prove its sensitivity in tracing regular differences.

While this setup is unrealistic, it is transparent, and allows us to control metrical principles independent of linguistic regularities. Figure B1 shows the resulting clustering of the “poems”: it is clear that the forms that employ caesurae can be distinguished from those that do not. The alignment works better for “syllabic” alexandrines because, unlike iambics, they have an additional distinctive feature – two fixed stressed positions versus an unregulated 12-syllable line.

Further tests confirm the outcome: for 100 runs with 20 poems generated for each form, we measure the cluster efficiency via Adjusted Rand Index (ARI). The median ARI for iambic meters is 0.55, indicating a somewhat noisy recognition of caesura-based forms. The median ARI for syllabics is 0.95 – an almost perfect clustering of the three varieties. Thus, we expect that the main structuring principle of caesura-based forms should be discoverable, even in very noisy settings.