Free Verse Prosodies: Identifying and Classifying Spoken Poetry Using Literary and Computational Perspectives (Rhythmicalizer)

Timo Baumann, Hussein Hussein, Burkhard Meyer-Sickendiek

Abstract *At least 80% of modern and postmodern poems exhibit neither rhyme nor metrical schemes such as iamb or trochee. However, does this mean that they are free of any rhythmical features? The US American research on free verse prosody claims the opposite: Modern poets like Whitman, the Imagists, the Beat poets and contemporary Slam poets have developed a post-metrical idea of prosody, using rhythmical features of everyday language, prose, and musical styles like Jazz or Hip Hop. It has spawned a large and complex variety in their poetic prosodies which, however, appear to be much harder to quantify and regularize than traditional patterns. In our project, we examine the largest portal for spoken poetry Lyrikline and analysed and classified such rhythmical patterns by using pattern recognition and classification techniques. We integrate a human-in-the-loop approach in which we interleave manual annotation with computational modelling and data-based analysis. Our results are integrated into the website of Lyrikline. Our follow-up project makes our research results available to a wider audience, in particular to high school-level teaching.*

Introduction

The theoretical starting point of our work is an American research discussion that has not previously been noted in German studies: the so-called free verse prosody. The most important theoretical orientation of this new theory of poetry was provided in 1980 by Charles O. Hartman in his influential study *Free Verse: An Essay on Prosody*. Hartman assumed that the free verse was prosodical, but not metrical—"the prosody of free verse is rhythmic organization by other than numerical modes"[1]—and therefore defined the prosody of the free verse via the term rhythm. Prosody was therefore a "system of rhythmic organization that governs the con-

1 Charles O. Hartmann: *Free Verse: An Essay on Prosody* (Princeton: University Press, 1980) 14.

struction and reading of a poem"[2]. Based on these theses, a discussion about the design principles of this free verse prosody has been developed in the US-American research area, reaching right down to the present day.

As an example, we offer the most important poem demonstrating the 'breath controlled line' in German poetry: Ernst Jandl's poem 'beschreibung eines gedichtes' (description of a poem) taken from the volume *der gelbe hund*:

bei geschlossenen lippen
ohne bewegung in mund und kehle
jedes einatmen und ausatmen
mit dem satz begleiten
langsam und ohne stimme gedacht
ich liebe dich
so daß jedes einziehen der luft durch die nase
sich deckt mit diesem satz
jedes ausstoßen der luft durch die nase
und das ruhige sich heben
und senken der brust[3]

The poem shown in its textual form is available as spoken by the author via the QR code printed above and can also be accessed via the URL https://www.lyrikline.org/de/gedichte/beschreibung-eines-gedichtes-1233. Even through reading the text one

2 Ibid.

3 Ernst Jandl: *der gelbe hund* (Darmstadt/Neuwied: Luchterhand, 1985).

will notice that the spoken delivery of the poem plays a vital function in its perceptive impression. It is this focus on *spoken poetry* that we focus on in our work.

Digitalization has made our field of research—the genre of audio poetry—possible in the first place. Of course, printed poems have been around for centuries, but audio poems have only existed at the present quantity and frequency since about 2000 with the development of the Internet. Reference has been made to the changed conditions of production and reception that prevail in the audio production of poetry in the digital age. While poetry readings change poems as text form in a lasting way, since they take place in the contingent space of the public in the age of the Internet, poetry can also be performed and intoned as a silent reading. According to Ursula Geitner, with the "end of public and domestic declamation practice around 1900", the "eloquence of the body is also reduced", insofar as this "new reading art […] concentrates on the author's voice."[4]

The idea of the project that we report in this chapter is twofold: we introduce the discussion on free verse prosody in post-modern poetry into German literary studies and examine the theses of this discussion. We also support them by using data-based methods not only in the literary analysis, but also in the computational models which we introduce and which leverage machine learning to foster our understanding of the relevant classes and their interrelation.

The remainder of the chapter is structured as follows: we introduce Rhythmicalizer, the project in which we have undertaken our research, in the following section and describe the literary theory foundation, the data and the methods that we have used. We describe our primary literary and computational results within the discussion of our methods, as these were produced in a cyclical, iterative manner by the interplay of literary analysis and computational modelling. We then discuss our overall learning that goes beyond the immediate results in our conclusion.

Project Rhythmicalizer

In this section, we describe the goal of the project for classiying poems according to their rhythmical features and the partners in the project.

The aim of the project *Rhythmicalizer* (www.rhythmicalizer.net) is twofold: we test and verify the basic patterns of free verse prosody theory, which in the US and for English is a highly recognized scientific theory but neglected as a lyrical rhythm for German, within the framework of a digital corpus analysis. It is complicated by the nature of free verse prosody which is an almost purely *spoken* aspect of a poem and is

4 Geitner, Ursula: *Die Sprache der Verstellung. Studien zum rhetorischen und anthropologischen Wissen im 17. und 18. Jahrhundert* (Berlin: De Gruyter, 1992), 342. Translation: B.M.S.

not readily observable from the pure textual form as in traditional metric schemes. We therefore base our analyses on the audible form of the poem (instead of, e.g., attempting to derive such features in a generic way from the written form). We then group poems into *classes* that express different kinds of free verse prosody.

In the computational part of the project, we integrate the two modalities—speech and text—and build an automatic classification system that uses quantitative features to yield our classifications. Our system needs to overcome the issue of *data sparsity* (i.e., the relatively small number of poems available as compared to the need for large amounts for data commonly used by modern machine-learning techniques).

Finally, we are concerned with the matter of preprocessing the poems in our collection so that they can be classified in an automatic way; and ultimately, we attempt to gather additional, new insight from the automatic classification by using a *human-in-the-loop* approach, in which we interleave manual annotation with computational modelling and data-based analysis.

Our research is performed in collaboration with the internet portal '*Lyrikline*' (www.lyrikline.org), located in Berlin and initiated by Literaturwerkstatt Berlin, probably the most important internet portal for international poetry readings worldwide hosting contemporary international poetry as audio (read by the authors) and text (original versions & translations). It provides the melodies, sounds and rhythms of international poetry, read by the original authors. Users can listen to the poet and read the poems both in their original language and in various translations.

An important aspect to the realization of the project was a change in German copyright laws in March 2018 regarding the use of publicly available data for research purposes.[5] It allowed us to freely download and use the data, but hardly did it simplify collaboration with *Lyrikline*.

Corpus Development

The database of our research project is based on the internet portal *Lyrikline*. The digital material on *Lyrikline* contains more than 10,000 poems by over 1200 international poets from almost 80 different countries. Nearly 80% of them are post-metrical poems. We examine German and English poems, with a total of 392 poets (227 German, 165 English) and about 3840 poems (2465 German, 1376 English). We

5 Burkhard Meyer-Sickendiek, Hussein Hussein: "The New German Copyright Law for Science and Education (UrhWissG): Consequences for DH-Projects Working with Non-Academic Partners", in: *Proceedings of the 15th Conference on Knowledge Organization WissOrg'17 of the German Chapter of the International Society for Knowledge Organization (ISKO) 2017*, Berlin, December 2017.

use this corpus to determine rhythmic patterns, extended with cornerstone poems by some authors that are not available on *Lyrikline* but on other platforms, such as YouTube.

As part of our project, we built a tool[6] to systematically extract poems and their meta-data from *Lyrikline* in order to form a corpus. We then performed forced text-audio alignment (which we refined manually as necessary) in order to relate speech and text line by line. In addition, we used various tools to extract syntactic information from the text as required (such as parts of speech for each word). We needed to label the data in order to detect the classes in them (rhythmical patterns). For this reason, the philologist classified every poem after listening to their audio recording into one of our rhythmical patterns.

Of course, a primary difficulty in a project such as ours is the expert categorization of poems into classes as not all poems on *Lyrikline* fit into a category and not every exceptional poem merits its own category. This was why we defined a threshold value for poems to constitute a new pattern: in total we expected up to 25 different patterns and used a threshold of 1% of poems belonging to this pattern. Thus, with about 2400 German-language poems on *Lyrikline*, the threshold value was 25: about 25 poem samples should be found per pattern for it to be relevant. (Note that not all classes reach this level; also, we could not manually fully analyze all of *Lyrikline*; see also Section 4.3.).

Table 1 shows some key descriptive statistics of the poems as assigned to their classes. The full manual categorization of poems into classes is available for research purposes from the authors.

Table 1: Description of the data used in the experiment.

	Poems	Lines	Characters	Audio
Lyrikline (German sub corpus)	2392	61849	2025484	52 hours
Cadence	31	1079	34163	49 min
Parlando	34	1435	44323	67 min
Variable foot	34	878	23684	39 min
Hook style	36	1090	33178	48 min
Gestic rhythm	33	897	27741	44 min
Ellipsis	56	2154	62704	104 min

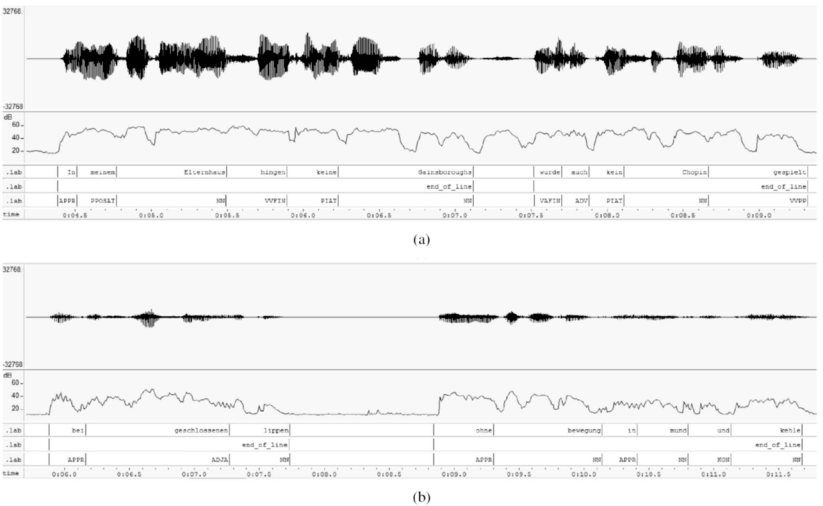
6 A program for extracting poems (text, audio and metadata) from www.lyrikline.org is available by request from the authors.

Permutation	30	1117	32041	46 min
Syllabic decomposition	21	540	12390	26 min
Lettristic decomposition	17	684	10007	31 min

Data Processing and Tools for Handling and Analysis

In order to simplify speedy manual analysis as well as hypothesis testing and to enable automatic processing, data on all poems are stored time-aligned, either in a word-by-word or in a line-by-line manner. Figure 1 shows an exemplary result of data preprocessing for two poems, each showing (from top to bottom) speech signal, intensity (dB), word alignment, end of line alignment, parser information (PoS-tagging), and time, by using the speech visualization tool Wavesurfer.[7]

Fig. 1: Poem analysis: (a) Analysis of the first two lines of the poem 'TEILS-TEILS' (English: Half Here, Half There) from the poet, Gottfried Benn, as an example of the 'Parlando' pattern. (b) As in (a) but for the poem 'beschreibung eines gedichtes' (English: Description of a Poem) from the poet, Ernst Jandl, as an example of the 'Variable Foot' pattern.



7 Kåre Sjölander, Jonas Beskow, "Wavesurfer–an open source speech tool", in: *Sixth International Conference on Spoken Language Processing*, 2000.

A forced alignment of text and speech for poems is performed using a text-speech aligner[8] which employs a variation of the Sail Align algorithm[9] implemented via Sphinx-4.[10] The line boundaries (the start of the first word and the end of the last word in each poetic line) are detected. The alignments are stored in a format that guarantees that the original text remains unchanged, which is important for recreating the exact white spacing in the poem (the white space is important as the text in a poem and helps readers know how to read a poem out). The forced alignment of text and audio in spoken poetry, especially in concrete and sound poetry, is non-trivial and often individual words or lines cannot be aligned. Therefore, the automatically extracted alignment information is manually corrected by the second author more than once (rectifying alignment information and in some cases correcting written text and audio files of poems) by listening to the audio file and looking at the waveform.

We also processed the text data of poems by using a statistical parser in order to extract syntactic features. The Stanford parser[11] is used to parse the written text of poems. The parser used the Stuttgart-Tübingen-Tag Set (STTS) table developed at the Institute for Natural Language Processing of the University of Stuttgart[12] for the parsing of German poems. A major problem in poem parsing is the absence of punctuation marks. In addition, many poems are written with special characters: sometimes the text is written in lower case with some words in upper case, which makes the recognition of sentence boundaries quite difficult. Furthermore, some sentences in the poems go beyond the line boundary and run on to the next line. Such unconnected syntactic elements result from the dissolution of poetic lines, caused by so-called enjambment. No countermeasure was taken by us to solve these problems since the correction of the problems was time-consuming and changed the poem form.

---

8 Timo, Baumann, Arne Köhn, Felix Hennig: "The Spoken Wikipedia Corpus Collection: Harvesting, alignment and an application to hyperlistening", *Language Resources and Evaluation*, vol. 52, no. 2, 303–329, 2018, special Issue representing significant contributions of LREC 2016.

9 Katsamanis, A., M. Black, P. G. Georgiou, L. Goldstein, and S. Narayanan, "Sail Align: Robust Long Speech-Text Alignment", in: *Proc. of Workshop on New Tools and Methods for Very-Large Scale Phonetics Research*, 2011.

10 Willie Walker et. Al., "Sphinx-4: A Flexible Open Source Framework for Speech Recognition", Mountain View, CA, USA, Tech. Rep., November 2004.

11 Anna Rafferty, Christopher D. Manning, "Parsing Three German Treebanks: Lexicalized and Unlexicalized Baselines", in: *Proceedings of the Workshop on Parsing German*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2008.

12 Anne Schiller et. al., "Guidelines für das Tagging deutscher Textcorpora mit STTS (Kleines und größes Tagset)", available on http://www.sfs.uni-tuebingen.de/resources/stts-1999.pdf, 1999 [accessed: 06.12.2019].

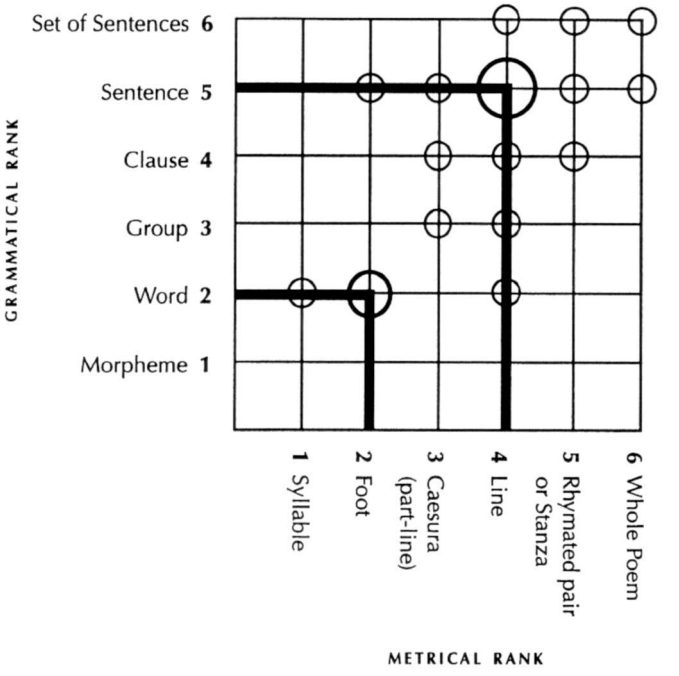
Method

This section describes the theoretical approaches used in the analysis and the computational approaches used for classification with the help of a web-based interface. It also describes the human-in-the-loop approach for classification and analysis of the corpus by using that interface.

Literary Analysis: Grammetrical Ranking and Rhythmic Phrasing

On the literary side, the classification is based on two theoretical approaches existing within the research field of free verse prosody: The 'grammetrical ranking'[13], and the 'rhythmic phrasing'[14].

Fig. 2: The vertical axis is the grammatical rank; the horizontal axis is the metrical rank. Intersection points help to identify the poem, for instance, the line arrangement (Wesling 1996).



13 Donald Wesling: *The Scissors of Meter: Grammetrics and Reading* (Ann Arbor, Michigan: University of Michigan Press, 1996).

14 Richard D. Cureton: *Rhythmic Phrasing in English Verse* (London: Longman, 1992).

The idea of grammetrical ranking has been developed by Donald Wesling, based on the key hypothesis that in poetry as a kind of versified language, the grammatical units (sentence, clause, group, word, morpheme) and the metrical units (syllable, foot, part-line, line, rhymated pair, stanza, whole poem) interact in a way for which Wesling finds 'scissoring' an apt metaphor (see Figure 1).

The idea of rhythmic phrasing was developed by Richard Cureton; it was based on Lerdahl and Jackendoff's 'Generative Theory of Tonal Music'.[15] Cureton divided the poetic rhythm into three components: metre, grouping and prolongation. Metre is about the perception of beats in regular patterns and grouping refers to the linguistic units gathered around a single climax or peak of prominence, quite similar to Wesling's ranking. Basically, Cureton's new idea is that of prolongation which refers to the anticipation and overshooting of a goal, the experience of anticipation and arrival.

In his groundbreaking study on 'The Prosodies of Free Verse' in 1971,[16] Donald Wesling distinguished five different prosodic types in the history of free verse prosody:

1. Whitmanic, referring to Whitman's adaptation of 'the biblical verset and syntax' in 'end-stopped lines .., with boundaries so often equivalent to those of larger units of grammar,' which Wesling sees as 'constitut[ing] the precomposition or matrix of free verse in English';
2. 'line-sentences,' as developed by Pound in Cathay on the basis of Ernest Fenollosa's theories of the sentence derived from study of Chinese;
3. dismemberment of the line, whereby the line becomes 'ground to the figures of its smaller units', and, as a sub-category, spatial dismemberment of the line by indentation, as by Williams in his triadic line verse;
4. systematic enjambment, whereby '[t]he lines.., are figures on the ground of the larger unit, the stanza'; and
5. dismemberment together with enjambment of the line, such that 'the middle units on the rank scale engage in a protean series of identity shifts as between figure and ground.'

Based on this typology extended by Wesling in later works,[17] we developed a fluency/disfluency spectrum based on nine different patterns. We illustrate this prosodic spectrum of fluency/disfluency by ranking these nine different poetic

15 Fred Lerdahl, Ray Jackendoff: *A Generative Theory of Tonal Music* (Cambridge, Mass: MIT Press, 1983).

16 Donald Wesling: "The Prosodies of Free Verse", in R. A. Brower (Ed.), *Twentieth-Century Literature in Retrospect* (Cambridge, Mass.: Harvard University Press, 1971).

17 Wesling, *Scissors of Meter*, 1996.

styles within the free verse spectrum. **(1) Cadence**: the cadence is the most fluent one. The basic idea of the cadence is the 'breath-controlled line' as an isochronous principle. Ezra Pound, who invented the idea of the cadence, was influenced by Chinese poetry that lacked any enjambments.[18] This explains the so-called line-sentence as the fundamental principle of the cadence. As different from the first class, more disfluent poems use 'weak enjambments' separating the nominal phrase and the verbal phrase of a sentence. Such 'weak enjambments' can be divided furthermore according to the emphases of the enjambments. **(2) Parlando**: the poems in the parlando style use 'weak enjambments' that do not emphasize the enjambments. **(3) Variable foot**: in contrast to parlando, the poems in the variable foot class emphasize the enjambments. These two classes are rather the fluent ones compared to those poems using 'strong enjambments', because the break in the reading is not really irritating as long as it is based on regular pauses in speech. Figure 1 shows an example of the rhythmical pattern 'variable foot' in the first two lines of a poem by the German poet Ernst Jandl: 'beschreibung eines gedichtes' (English: Description of a Poem). As it is obvious from the intensity or power contour, the poet makes a short stop after each colon in the sentence, imitating the regular pauses in speech by taking into account his breathing as a rhythmical feature (pause between words 'lippen' and 'ohne'). We used a text-speech aligner for the detection of word and sentence boundaries.

A more disfluent kind of poetry uses strong enjambments, which means it separates articles or adjectives from their nouns or even splits a word across lines, as in Paul Celan's poems. Poems using 'strong enjambments' can also be divided according to the emphases of the enjambments. **(4) Hook style** (German: Hakenstil): the poems in the hook style use 'strong enjambments' that do not emphasize the enjambments. **(5) Gestic rhythm**: in comparison to the hook style, the poems in the gestic rhythm emphasize the enjambments, meaning that the author makes an irritating break after each line when reading his poem, although the sentence should continue. This fifth pattern was invented by Bertolt Brecht and used in his later works. It had a huge impact on poets from the former German Democratic Republic.

Moving towards the radical disfluent pole, the next pattern is the ellipsis. **(6) Ellipsis**: it is the omission of one or more grammatically necessary phrases. This rhetorical figure can also affect the prosody of a poem, which has been observed in poems of Paul Celan. **(7) Permutation**: it is a conversion or exchange of words or parts of sentences or a progressive combination and rearrangement of linguistic-semantic elements of a poem, a principle that was very popular in German 'concrete poetry'.[19] Even more radical kinds of poetic disfluency have been developed in mod-

18 George Steiner: *After Babel—Aspects of Language and Translation* (Oxford: Oxford University Press, 1975), 358.

19 Ulrich Ernst: „Permutation als Prinzip in der Lyrik". *Poetica*, vol. 24, no. 3/4, 225–269, 1992.

ern 'sound poetry' by dadaistic poets like Hugo Ball and Schwitters or concrete poets like Ernst Jandl. Within the genre of sound poetry, there are two main patterns: **(8) Syllabic decomposition**, dividing the words into syllables; and **(9) Lettristic decomposition**, the last and most disfluent pattern, which divides the words into single characters, as for example in Ernst Jandl's famous poem 'schtzngrmm'.

Computational Analysis

We devised our computational methods with the primary concerns on *interpretability* (i.e., the ability of a method to explain what aspects of the data determine its behaviour) and the ability to deal with *data sparsity* (i.e., the limitation that there is too little data to fully train advanced machine learning algorithms).

Poetic Prosody Classification

Our project setup allows *supervised learning* where the parameters of a computational model are tuned to data (in our case poems) that have already been classified into a range of classes. The goal of the classifier is to extract those aspects of a poem that characterize it as part of a class, while neglecting other aspects, so that the resulting characterizations generalize to previously unseen poems.

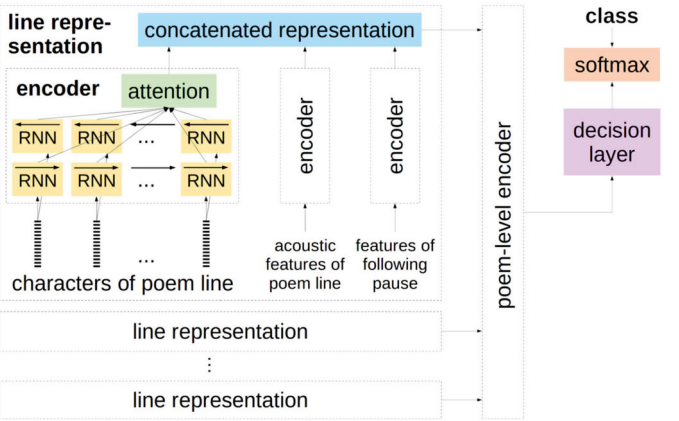
We first focused our efforts on building interpretable classifiers, such as decision trees, and on specific interpretable features, such as the presence or absence of a verb in a line.[20] Obviously, most poems consist of many lines and, therefore, the question of how to integrate knowledge across lines arises (e.g., what should be the proportion of lines that exhibit a certain feature to be deemed relevant?). Of course, there is no clear answer and the question of what aggregation of what features in a line yields the best performance assumes some significance (add to it aspects such as *speaking style* which cannot easily be measured by a discrete feature). We therefore focused more and more on deep-learning-based methods that help to discover latent or hidden structure in data across numerous dimensions (such as in the text, the speech, and other aspects of recitation). Particularly, based on the literary theory, we determined that pauses between lines were a critical aspect in addition to the speech and the text.

Deep learning is known to require a great amount of data and although the number of poems manually analysed in the course of the project may appear large, com-

20 Hussein Hussein, Burkhard Meyer-Sickendiek, Timo Baumann: "Tonality in Language: The 'Generative Theory of Tonal Music' as a Framework for Prosodic Analysis of Poetry", in: *Proceedings of the 6th International Symposium on Tonal Aspects of Languages (TAL) 2018*, Berlin, Germany, June 2018a; Hussein Hussein, Burkhard Meyer-Sickendiek Timo Baumann, "Automatic Detection of Enjambment in German Readout Poetry", in: *Proceedings of Speech Prosody*. Poznán, Poland, 2018b.

plex machine-learning models would require many orders of magnitude more data to be able to work at their full potential.

Fig. 3: Full model for deep-learning-based poetry style detection: each line is encoded character-by-character by a recurrent neural network (using GRU cells) with attention. Acoustic features of each line and the pause following up to the next line are encoded similarly. Per-line representations are concatenated and passed to a poem-level encoder. The final decision layer optimizes for the poem's class.



Also, the fact that the prosodic structure of a poem is reflected in many (or most) lines of a poem that belongs to a certain pattern, can be exploited by *hierarchical attention models*[21] which presuppose structure in a given document.[22] In our case, we know that a poem consists of lines and given text-speech alignment, we also know both the text and the speech audio that correspond to the line. Furthermore, given that the pause at the end of a line is stylistically relevant, for example, to differentiate different types of enjambment, we also encode the pause that separates a line from the next. We combine the information from these three aspects of a line (text, speech and following pause) using three recurrent neural networks (bidirectional GRUs) that furthermore use inner-attention (also called self-attention) so as to focus on what is most relevant in the line. In order to combine the information from all

21 Zichao Yang et al.: "Hierarchical attention networks for document classification", in: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics*: Human Language Technologies. pp. 1480–1489, Association for Computational Linguistics, 2016.

22 In this case, "document" refers to a poem.

lines into our class assignment, we use yet another recurrent layer across the lines. This architecture is depicted in Figure 3. As a result, our classifier is able to single out the information that differentiates stylistic patterns as it is trained to be optimal towards differentiating the prosodic classes in the training material.

Given the fact that most of the lines exhibit the structural properties that yield a poem's prosodic classification, we use a two-stage training procedure, in which we first train the line-by-line classification in isolation and only afterwards do we address the full network including the final decision layer. In this way, we make most efficient use of our data and successfully tackle the issue of data sparsity. We furthermore improve results by pretraining part of our representation encoders using poetic texts (from www.deutschestextarchiv.de) and general speech audio. We set aside testing material and find that our classifiers achieve high performance.

After establishing that deep learning could be used to classify lines of poems according to poetic prosodic style,[23] we extended it to experiments on classifying poetry by using a hierarchical attention network.[24] We found that free verse prosodies could be classified along a fluency continuum and that the classifier's wrong classifications cluster along this continuum. Through ablation studies (i.e., leaving out certain features to find their relative importance for the final results), we found that, depending on the classes analysed, the acoustic realization of the speech itself is highly relevant for classification, so is the realization of the pause following each line of the poem. In fact, our classifier is highly reliable for determining different kinds of enjambments and an additional annotation of enjambments is unnecessary in the poems (unlike what we had originally hypothesized). We also tried traditional classification approaches by using engineered features and, in a comparison with our deep-learning-based method, found that they did not work as well although they had direct access on theoretically grounded features.[25] We also found that features of a poetic line (such as enjambment) could be identified by a neural network and that their annotation did not improve the overall poem classification.[26] Finally, we integrate engineered features into the deep-learning approach which further boosts

23 Burkhard Meyer-Sickendiek, Hussein Hussein, Timo Baumann: "Recognizing Modern Sound Poetry with LSTM Networks", in: *Proceedings of Elektronische Sprachsignalverarbeitung (ESSV)*. TUDpress, 192–199, 2018.

24 Timo Baumann, Hussein Hussein, Burkhard Meyer-Sickendiek: "Style Detection for Free Verse Poetry from Text and Speech", in: *Proceedings of COLING*. Santa Fe, USA, 1929–1940, 2018a.

25 Timo Baumann, Hussein Hussein, Burkhard Meyer-Sickendiek: "Analysis of Rhythmic Phrasing: Feature Engineering vs. Representation Learning for Classifying Readout Poetry", in: *Proceedings of the Joint LaTeCH&CLfL Workshop*. Santa Fe, USA, 44–49, 2018b.

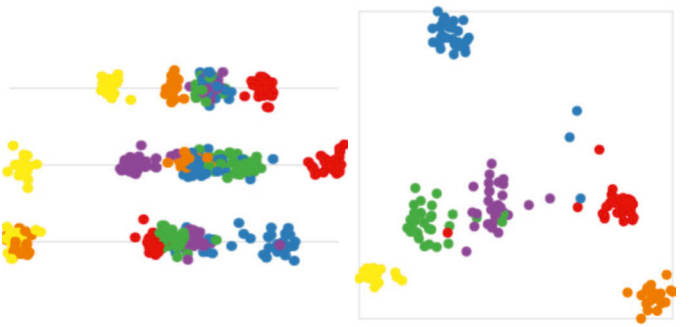
26 Timo Baumann, Hussein Hussein, Burkhard Meyer-Sickendiek: "Analysing the Focus of a Hierarchical Attention Network: The Importance of Enjambments When Classifying Postmodern Poetry", in: *Proceedings of Interspeech*. Hyderabad, India, 2162–2166, 2018c.

performance.[27] We refer the interested reader to our individual publications for details on the methods used.

Visualization of Results

Our results can provide multiple kinds of visualization which may help literary analysis. Visualization is particularly useful for deep-learning-based approaches as it provides some understanding of the inner workings of the model that is otherwise hard to understand. For example, representations of a poem (to be used for classification) can be forced into lower dimensions and the placement of poems in these dimensions can be visualized in 1D, 2D or 3D[28] spaces, as some are shown in Figure 4. The analysis of outliers (e.g., the one red poem among many green poems) has proved to be useful for literary analysis.

Fig. 4: Visualization of low-dimensional mappings of the poems contained in six classes (Left: Results depend on random initialization and three results are shown).

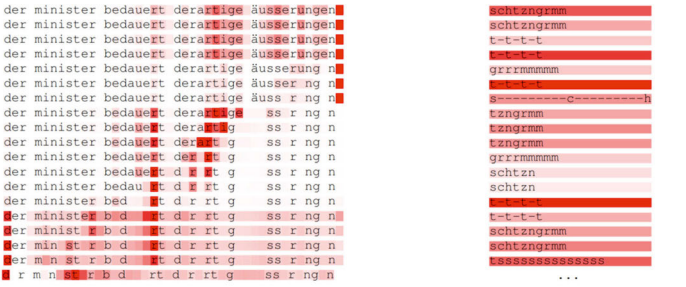


The attention mechanism in our model determines what the model pays attention to in a certain poem and this can be exploited for visualization. However, it needs to be noted that the attention is primarily focused on those aspects that allow the model to differentiate classes, not to describe classes. It can nonetheless shed light on classifier decisions, as can be seen in Figure 5.

27 Hussein Hussein, Burkhard Meyer-Sickendiek, Timo Baumann: "Free Verse and Beyond: How to Classify Post-modern Spoken Poetry", in: *Proceedings of Speech Prosody*. Tokyo, Japan, 690–694, 2020.

28 An interactive version of a 3D plot of poems analysed in this way is available at https://web.archive.org/web/20210404220806/https://timobaumann.bitbucket.io/colingfreeversepoetry/.

Fig. 5: Visualization of attention in two lettristic poems: (a) attention to characters within the line, (b) attention to lines (including the audio) in the poem.



(a) Jaap Blonk: Der Minister I

(b) Ernst Jandl: schtzngrmm

Finally, a different form of visualization concerns the presentation of our subject matter—the free verse prosody—by using data from our partner *Lyrikline*. Meanwhile *Lyrikline* has visualized the classes we investigated on the website of *Lyrikline*, so an own link to *Lyrikline* represents our classes. This form of visualization is of course so important that our results can now be used for teaching poetics and prosody.

Iterative Human-in-the-Loop Approach

Above, we have described the methods we used for literary and computational analysis, respectively. We have selected poems from *Lyrikline* that form one of the multiple patterns and we have built computational models that are able to classify poems based on these patterns with relatively high performance. However, we have already mentioned the issue of data sparsity and it is tedious and prone to oversight errors in manually selecting poems that belong to a pattern. We therefore come back to our original ideas[29] and combine both parts of the project to iteratively analyse many more poems in the *Lyrikline* collection. We have implemented an interface for *humanist-in-the-loop* classification and analysis[30] that allows a literary scholar to manually assign poems that have been automatically classified to a certain class. We have gradually applied this method to cover the full corpus.

29 Timo Baumann, Burkhard Meyer-Sickendiek: "Large-scale Analysis of Spoken Free-verse Poetry", in: *Proceedings of Language Technology Resources and Tools for Digital Humanities (LT4DH)*, 125–130, Osaka, Japan, December 2016.

30 Timo Baumann et al.: "A Tool for Human-in-the-Loop Analysis and Exploration of (not only) Prosodic Classifications for Post-modern Poetry", in: *Proceedings of INF-DH*. Kassel, Germany. Gesellschaft für Informatik, 151–156, 2019.

The main contribution is the integration of a human-in-the-loop process via a web-based interface. The web interface uses the results of the automatic classifier and makes as much information about the decision making process available as is possible. In so doing, the interface also includes the model's confidence (the trust of the model in its judgment) and its attention (places in the text and speech that the model deems particularly relevant for the judgment). The interface enables a philologist to explore the classification results and potentially correct or enhance them.

During the analysis of our data we repeatedly find a pattern that is not yet defined and we include it in our corpus. This pattern must again go through the quantification process. If there are more than 25 examples of this new pattern, then it is a relevant pattern of modern and postmodern free verse prosody. That is, we assume a threshold value that helps us to compensate for the difference between qualitative and quantitative analysis.

Our interface for poem classification allows us to add new classes and change prior class assignments in the light of new classes. For example, it allows splitting a class into multiple sub-classes and then retraining the computational models as the exploration of the corpus proceeds in order to help with further analyses.

In the end and over the course of the project, especially in the literary studies part of the project, a total of 18 rhythmic patterns have been developed: Longline Poems; Parlando; Flows; Free Associations; Cadence; Unstressed Enjambments; Variable Foot; Sprungrhythm; Syncopations; Rubato; Permutation; Gestic Rhythm (= Stressed Enjambments); Cut-ups; Dialect Poems; Ellipses; Staccato; Syllabic Decomposition; and Lettristic Decomposition. All these patterns are explained and illustrated by a number of examples on the *Lyrikline* website. Also, a monograph furthermore explains the theoretical framework, the patterns mentioned and the development of free verse prosody in modern and postmodern poetry.[31]

Conclusion

To conclude, we first summarize our approach and discuss the outcome of our project before we briefly describe some future work.

In the Rhythmicalizer project we have melded computational and literary methods to introduce ideas of free verse poetry into the debate on present-day post-modern German poetry. We have linked the ideas of grammetrical ranking as well as rhythmic phrasing and found a fluency/disfluency continuum along which many poems can be classified.

31 Burkhard Meyer-Sickendiek: *Hörlyrik. Eine interaktive Gattungstheorie* (München/Paderborn: Brill/Fink, 2020).

In the computational part of the project, we have shown that it is possible to build classifiers based on speech and text that are able to differentiate the prosodic patterns despite the high complexity and sparsity of the given poems. The resulting classifiers show patterns that underline the fluency/disfluency continuum.

We have finally built a *humanist-in-the-loop* interface that enables us to further analyse and classify a large portion of the poems available on *Lyrikline by* using a data-driven methodology to work (read and listen) through thousands of poems.

In our project, we were able to analyse a large portion of the German poems on *Lyrikline* and to trace the development of free-verse prosody in modern and postmodern poetry.

A central limitation of our approach is the reliance on the differentiation of classes. There are always uncertainties in classification, in particular for something that is as stylistically diverse and continuous as poetry. They primarily consist in the polyvalence of poems which can never really be assigned to a single class or characteristic. On the contrary, in good poems, several patterns or characteristics can often be found, which is why the assignment to exactly one pattern is always fraught with uncertainties. In this case, it is important to add considerable weight to the most important characteristics from a literary historical point of view. To give an example: In expressionist poems, the so-called 'Reihenstil' (=sequence style) is the most important feature, that is, the sequence of events that are heterogeneous in themselves. There are, however, other characteristics that characterize an expressionist poem. This issue could have been overcome by introducing a hierarchy of classes or by establishing a multi-class labelling approach. While potentially feasible, such approaches complicate the computational analysis and only partially alleviate the problem. We have therefore limited the analysis to plain classification.

Any classification of a work of art is always an abstraction, neglect of details. We could only classify the poems on *Lyrikline* by following a strategy of simplification and reduction. This strategy became even more extreme when it came to the characteristics of computational learning, as these characteristics were necessarily simpler and more reduced than the philologist's observations. At this point the calibration or threshold value is again important: Is the detailed observation possibly not a single case but a recurring pattern?

Another limitation is the type of machine learning used for classification. A neural network (and most other classification algorithms) does not model the classes and their properties, but merely differentiates between classes. Therefore, any analysis of the learned classifiers is limited and differs significantly from a human perspective which aims at forming classes based on their inner coherence. The hermeneutical idea of learning, in contrast, is way more based on the idea that different 'horizons' of human experiences can be mingled during the learning phase. These 'horizons' obviously do not exist in machine learning. Here we refer to Hans Georg Gadamer's concept of 'merging horizons', which describes the process

of understanding as a mixture of classical texts and works of art with a contemporary horizon, that is, a process of understanding between different cultures. This form of learning is naturally missing in artificial intelligence.

The second central conflict concerns the accentuation of our classification by the provider *Lyrikline* or by the authors of *Lyrikline* classified by us. It is not always acceptable for lyricists or artists to see their work classified and assigned in a certain way. As a rule, this may only be done with artists or writers who died long ago. Living artists often refuse to be classified. We hope (and some positive examples underpin this) that our data-driven endeavour nonetheless convinces them that classification of their work does not diminish or demolish their work.

Over the course of the project, multi-modal processing has become more commonplace (and the tools that can freely be used are easier to use) in the Digital Humanities, and deep-learning-based approaches are also more common. We hope that our project sets a good example for the integration of multi-modal data with an iterative human-in-the-loop approach to data analysis that yields interesting problems and results both in computational studies and in the humanities.

At present, we are involved in transferring our research results to humanities teaching at the high school level. Our interactive web-based human-in-the-loop project can be used by students to (a) learn about present-day poetry and the various forms that it takes, to (b) reflect about language use in poetry, for example, to learn how to differentiate between enjambments based on their use in a poem by singling out and paying attention to details, and to (c) understand how a topic as liquid and hard to grasp as poetry can nonetheless be differentiated, categorized, classified or hierarchized. These competences go far beyond their immediate use in poetry analysis and are highly relevant for future humanities scholars in a computational age.

Bibliography

Baumann, Timo, Arne Köhn, and Felix Hennig: "The Spoken Wikipedia Corpus Collection: Harvesting, alignment and an application to hyperlistening", *Language Resources and Evaluation*, vol. 52, no. 2, 303–329, 2018, special Issue representing significant contributions of LREC 2016.

Baumann, Timo, Hussein Hussein, and Burkhard Meyer-Sickendiek: "Style Detection for Free Verse Poetry from Text and Speech", in: *Proceedings of COLING*. Santa Fe, USA, 1929–1940, 2018a.

Baumann, Timo, Hussein Hussein, and Burkhard Meyer-Sickendiek: "Analysis of Rhythmic Phrasing: Feature Engineering vs. Representation Learning for Classifying Readout Poetry", in: *Proceedings of the Joint LaTeCH&CLfL Workshop*. Santa Fe, USA, 44–49, 2018b.

Baumann, Timo, Hussein Hussein, and Burkhard Meyer-Sickendiek: "Analysing the Focus of a Hierarchical Attention Network: The Importance of Enjambments When Classifying Post-modern Poetry", in: *Proceedings of Interspeech*. Hyderabad, India, 2162–2166, 2018c.

Baumann, Timo, Hussein Hussein, Burkhard Meyer-Sickendiek, and Jasper Elbeshausen: "A Tool for Human-in-the-Loop Analysis and Exploration of (not only) Prosodic Classifications for Post-modern Poetry", in: *Proceedings of INF-DH*. Kassel, Germany. Gesellschaft für Informatik, 151–156, 2019.

Baumann, Timo and Burkhard Meyer-Sickendiek: "Large-scale Analysis of Spoken Free-verse Poetry", in: *Proceedings of Language Technology Resources and Tools for Digital Humanities (LT4DH)*, 125–130, Osaka, Japan, December 2016.

Cho, Kyunghyun, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio: "Learning phrase representations using rnn encoder–decoder for statistical machine translation", in: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1724–1734, Doha, Qatar. Association for Computational Linguistics, 2014.

Cureton, Richard D.: *Rhythmic Phrasing in English Verse*. Longman, 1992.

Ernst, Ulrich.: "Permutation als Prinzip in der Lyrik". *Poetica*, vol. 24, no. 3/4, 225–269, 1992.

Geitner, Ursula: *Die Sprache der Verstellung. Studien zum rhetorischen und anthropologischen Wissen im 17. und 18. Jahrhundert*, De Gruyter, 1992.

Hartman, Charles O.: *Free Verse: An Essay on Prosody*. Princeton University Press, 1980.

Hussein, Hussein., Burkhard Meyer-Sickendiek, and Timo Baumann: "Tonality in Language: The "Generative Theory of Tonal Music" as a Framework for Prosodic Analysis of Poetry", in: *Proceedings of the 6th International Symposium on Tonal Aspects of Languages (TAL) 2018*, Berlin, Germany, June 2018a.

Hussein, Hussein, Burkhard Meyer-Sickendiek, and Timo Baumann, "Automatic Detection of Enjambment in German Readout Poetry", in: *Proceedings of Speech Prosody*. Poznán, Poland, 2018b.

Hussein Hussein, Burkhard Meyer-Sickendiek, and Timo Baumann: "Free Verse and Beyond: How to Classify Post-modern Spoken Poetry", in: *Proceedings of Speech Prosody*. Tokyo, Japan, 690–694, 2020.

Jandl, Ernst: *der gelbe hund*, Darmstadt und Neuwied: Luchterhand, 1985.

Katsamanis, Athanasios, Matthew Black, Panayiotis G. Georgiou, Louis Goldstein, and Shrikanth S. Narayanan, "Sail Align: Robust Long Speech-Text Alignment", in: *Proc. of Workshop on New Tools and Methods for Very-Large Scale Phonetics Research*, 2011.

Lerdahl, Fred and Ray Jackendoff: *A Generative Theory of Tonal Music*. Cambridge, Mass. MIT Press, 1983.

Meyer-Sickendiek, Burkhard.: *Hörlyrik. Eine interaktive Gattungstheorie*. Brill/Fink, Paderborn, München 2020.

Meyer-Sickendiek, Burkhard andHussein Hussein: "The New German Copyright Law for Science and Education (UrhWissG): Consequences for DH-Projects Working with Non-Academic Partners", in: *Proceedings of the 15th Conference on Knowledge Organization WissOrg'17 of the German Chapter of the International Society for Knowledge Organization (ISKO)* 2017, Berlin, December 2017.

Meyer-Sickendiek, Burkhard, Hussein Hussein, and Timo Baumann: "Recognizing Modern Sound Poetry with LSTM Networks", in: *Proceedings of Elektronische Sprachsignalverarbeitung (ESSV)*. TUDpress, 192–199, 2018.

Rafferty, Anne N. and Christopher. D. Manning, "Parsing Three German Treebanks: Lexicalized and Unlexicalized Baselines", in: *Proceedings of the Workshop on Parsing German*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2008, 40–46.

Schiller, Anne, Simone Teufel, Christine Stöcker, and Christine Thielen, "Guidelines für das Tagging deutscher Textcorpora mit STTS (Kleines und größes Tagset)", Available on http://www.sfs.uni-tuebingen.de/resources/stts-1999.pdf, 1999, last accessed on 06. December 2019.

Sjölander, Kånder and Jonas Beskow, "Wavesurfer–an open source speech tool", in: *Sixth International Conference on Spoken Language Processing*, 2000.

Steiner, George: *After Babel—Aspects of Language and Translation*. Oxford University Press, 1975.

Walker, Willie, Paul Lamere, Philip. Kwok, Biksha Raj, Rita Singh, Evandro Gouvea, Peter Wolf, and Joe Woelfel, "Sphinx-4: A Flexible Open Source Framework for Speech Recognition", Mountain View, CA, USA, Tech. Rep., November 2004.

Wesling, Donald: "The Prosodies of Free Verse", in *Twentieth-Century Literature in Retrospect*. R. A. Brower (Ed.), Harvard University Press, 1971.

Wesling, Donald: *The Scissors of Meter: Grammetrics and Reading*. University of Michigan Press, 1996.

Yang, Zichao, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy: "Hierarchical attention networks for document classification", in: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. pp. 1480–1489, Association for Computational Linguistics, 2016.