TAL2018, Sixth International Symposium on Tonal Aspects of Languages
18-20 June 2018, Berlin, Germany

Tonality in Language: The "Generative Theory of Tonal Music" as a Framework for Prosodic Analysis of Poetry

Hussein Hussein[1], *Burkhard Meyer-Sickendiek*[1], *Timo Baumann*[2]

[1]Department of Literary Studies, Free University of Berlin, D-14195 Berlin, Germany
[2]Language Technologies Institute, Carnegie Mellon University, Pittsburgh, USA
{hussein, bumesi}@zedat.fu-berlin.de, tbaumann@cs.cmu.edu

Abstract

This contribution focuses on structural similarities between tonality and cadences in music on the one hand, and rhythmical patterns in poetic languages respectively poetry on the other hand. We investigate two exemplary rhythmical patterns in modern and postmodern poetry to detect these tonality-like features in poetic language: The *Parlando* and the *Variable Foot*. German poems readout from the original poets are collected from the webpage of our partner *lyrikline*. We compared these rhythmical features with tonality rules, explained in two important theoretical volumes: The *Generative Theory of Tonal Music* and the *Rhythmic Phrasing in English Verse*. Using both volumes, we focused on a certain combination of four different features: The grouping structure, the metrical structure, the time-span-variation and the prolongation, in order to detect the two important rhythmical patterns which use tonality-like features in poetic language (*Parlando* and *Variable Foot*). Different features including pause and parser information are used in this classification process. The best classification result, calculated by the f-measure, for *Parlando* and *Variable Foot* is 0.69.

Index Terms: tonality in poetic language, prosodic patterns in postwar poems, automatic detection of prosodic features

1. Introduction

Tonality is an organized system of tones (e.g., the tones of a major or minor scale) in which one tone (the tonic) becomes the central point for the remaining tones. The tonic is the tone of complete relaxation, the target toward which other tones lead. This relaxation is caused by the cadence: a musical chord sequence moving to a harmonic close or point of rest. The *Generative Theory of Tonal Music* (GTTM), conceived by the music theorist Fred Lerdahl and the linguist Ray Jackendoff [1], is certainly the most important attempt to detect and to analyze the structural similarities between tonality in music and grammar in language. Deeply influenced by Noam Chomsky's transformational or generative grammar, Lerdahl and Jackendoff developed a musical grammar based on similar tree structure-style hierarchical organizations uniting musical "phrase groupings", or simply groupings. Such a grouping distinguishes the notion of phrases as relatively closed, or self-contained, musical units from that of the articulated phrasing associated with performance. An example of such a group is the musical phrase: "the smallest musical unit that conveys a more or less complete musical thought. Phrases vary in length and are terminated at a point of full or partial repose, which is called a cadence." [2] pp. 43-44.

1.1. Phrase Groupings in Poetry

Such phrases and cadences as examples for tonal structures could also be found in language, especially in poetic language. Eleanor Berry [3] offered a number of examples for such groupings to be found in modern American poetry, all of them representing typical line arrangements in modern poems: 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 Donald 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, in turn derived from the 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 the poet William Carlos Williams does in his triadic line verse; 4) systematic enjambment, whereby the lines are "figures on the ground of the larger unit, the stanza"; 5) dismemberment 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" [3] pp. 880.

This paper focuses on the third of these five patterns, the "triadic line", also known as the *Variable Foot*. Williams developed this famous *Variable Foot* in his late volumes *The Desert Music* (1954), *Journey to Love* (1955), and *Paterson V* (1958). The *Variable Foot* is based on the idea that, despite the different number of syllables per line, all the lines are isochronic, because all lines are based on a similar phrase/clause. In his readings, Williams emphasized the isochronicity of the lines by interrupting each by a regular breathing pause, so that each line seems to be based on a single breath unit. The following poem of the poet Ernst Jandl - *Beschreibung eines Gedichts* [4] pp. 129 - uses exactly the same rhythmic pattern, the *Variable Foot*:

"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
das ruhige sich heben
und senken der brust"

Jandls uses the *Variable Foot* and its "breath-controlled line" [5] pp. 83, which divides the syntax into a phrase or clause per line. That each line corresponds to exactly one single breath unit, causing a short break - a breathing space - at the end of each line, becomes obvious with regards to Figure 1b: There is a characteristic gap at the end of the first line.

With regards to similar "phrase groupings" in modern and postmodern poetry, we will compare the *Variable Foot* with a similar pattern, also using this "sub-category below the sentence-level", that is a phrase/clause in each line. This second rhythmi-

10.21437/TAL.2018-36

cal pattern is called the *Parlando*, which was also very common in postwar German poetry. It was developed by the German poet Gottfried Benn. The *Parlando* is a prosodic style similar to the litany, using a similar orientation towards everyday speech in order to express the speaker's spontaneous feelings. The most famous example is Benns poem "Teils-Teils" [6] pp. 317:

"In meinem Elternhaus hingen keine Gainsboroughs
wurde auch kein Chopin gespielt
ganz amusisches Gedankenleben
mein Vater war einmal im Theater gewesen
Anfang des Jahrhunderts
Wildenbruchs »Haubenlerche«
davon zehrten wir
das war alles."

Both these patterns – the *Variable Foot* as well as the *Parlando* - belong to the third example of Berrys five patterns using a similar kind of line-arrangement. The important difference lies in the phrasing: The *Parlando* makes no use of the "breath-controlled line" [5] pp. 83. Just like the *Parlando*, the *Variable Foot* had a huge impact on German poetry beginning in the same period, the 1960s and 1970s. For this reason, we believe that many examples of both patterns can be found in our corpus, the modern and postmodern poems in the *lyrikline* database. It is our project's contention that the dominant grouping structures in musical tonality - the phrase and the cadence - can also be found in the prosodic patterns used in modern and postmodern poems, especially those two patterns mentioned: the *Parlando* and the *Variable Foot*. Both these patterns have a tonality-like structure using a syntactic sequence moving to a harmonic close or point of rest. The exemplary analysis is particularly devoted to the GTTM, respectively to Richard Cureton's theory of *Rhythmic Phrasing in English Verse* (RPEV) [7] which is based on the GTTM. The GTTM and the RPEV both offer a very fruitful framework for the manual and digital analysis of these rhythmic patterns and for the specific "tonality" of (post-) modern poems.

1.2. Applying the GTTM to Poetry Analysis

A similar attempt to compare the dominant grouping structures in musical tonality and in *free verse prosody* already took place in the 1980s by two important theoretical approaches. The first one already mentioned is the *Generative Theory of Tonal Music*. This theory is based on four hierarchical systems that shape our musical intuitions: 1) The *Grouping structure* is based on the hierarchical segmentation of the musical piece into motives and phrases. 2) The *Metrical structure* identifies the regular alternation of strong and weak beats at a number of hierarchical levels, differing between the beat and the time span between two beats. Both two structures explain the so-called "time-span segmentation". 3) The *Time-span reduction* combines the information gleaned from these metrical and grouping structures. This is illustrated in a tree structure-style hierarchical organization uniting time-spans at all temporal levels. 4) The *Prolongational reduction* provides our "psychological" awareness of tensing and relaxing patterns in a given musical piece: In a strong prolongation, the roots, bass notes, and melodic notes are identical which effects the feeling of continuity and progression, caused by a movement towards relaxation.

Given this theory, both poetic patterns – the *Variable Foot* as well as the *Parlando* – offer a similar line arrangement and a similar kind of prolongation, caused by the incomplete syntax at the end of nearly each line: the meaning runs over from one poetic line to the next. But in the *Parlando*-style the poet does not emphasize the stops at the end of each line, in difference to those poets using the *Variable Foot*-pattern. This is obvious when listening to the audio recordings of both patterns.

To detect this difference with regards to the poems on *lyrikline*, we make use of Richard Curetons theory on RPEV. Cureton proved that the phrasing is a typical pattern not just in tonal music, but also in poetry. Following the hierarchical system by Lerdahl and Jackendorff, Richard Cureton has divided the poetic rhythm into three (not four) components: meter, grouping and prolongation [7] pp. 124. The meter contains the perception of beats in regular patterns, the grouping refers to the linguistic units gathered around a single climax of prominence, and the prolongation refers to the anticipation and overshooting of a goal, such as the end of a line in an enjambment. Cureton's rhythm theory involves the interrelationship of these three components within a strictly hierarchical structure. A rhythm consists of a series of local events or units that are perceived as more or less prominent elements within longer events or units, which in turn are perceived as more or less prominent elements within even longer events or units, and so on to the entire poem. The analysis of phrase movements for Cureton involves examining the interaction of grouping and prolongation in a hierarchical organization. With regards to the both patterns – *Parlando* as well as *Variable Foot* – Cureton offered a new insight by "defining these line-terminal syntactic expectations as mid-level prolongational energies" [7] pp. 153. That means: Both patterns involve the experience of anticipating a goal at the end of each line, caused by the enjambment and its connection to the second part of the sentence in the following line. So both patterns use the prolongation in nearly every line. But only the *Parlando* ignores this prolongation and its enjambment by arriving immediately at the goal in the next line. In other words: Only in the *Parlando*, the authors reading includes a time-span-reduction.

1.3. Research Question and Hypothesis

We focus on structural similarities between tonality and cadences in music as well as poetic languages by using hermeneutical and computational methods. The aim is to detect the tonality-like features of both rhythmical patterns (*Parlando* and *Variable Foot*) in a corpus of modern readout poetry by using automatic classification methods.

The paper is organized as follows: Section 2 reviews the philological method and digital tools used in the analysis as well as the features utilized in the classification process. The experimental results are described in Section 3. Finally, conclusions and future works are presented in Section 4.

2. Method

This section describes the corpus and tools as well as features and classifiers used in the experiment.

2.1. Database

In the project *Rhythmicalizer* (`www.rhythmicalizer.net`), we want to offer a theoretical as well as digital framework for the automatic recognition of rhythmical patterns in modern and postmodern poetry. Therefore, we use the database of our partner *lyrikline* (`www.lyrikline.org`), which contains speech and text data of modern and contemporary poetry, giving us access to hundreds of hours of author-spoken poetry. *Lyrikline* hosts contemporary international poetry as audio files (read by the authors themselves) and texts (original versions & translations). The digital material covers more than $10,800$ poems by more than $1,200$ international poets from 80 different

languages. Nearly 80% of the *lyrikline*-poems are postmetrical poems. In this project, we will use all poems written in English and German (more than 3,600 poems). The total number of poets writing in German and English is 215 and 154, respectively.

The philological scholar (second author) in the project collected the poems written in German from the *lyrikline* website, whereas he just go through all the poems by hand and label them as *Parlando* or *Variable Foot* based on his experience. The total number of poems in this study is 68 from 24 poets (34 poems in each class, i.e. *Parlando* and *Variable Foot*). This selection is about 11% of all the poets speaking German. Each of these authors read at least one and at most nine poems. The minimal and maximal number of lines in poems is 9 and 231, respectively. The length of audio files is between 27 and 715 seconds.

2.2. Processing Tools

The first step for the identification of poems as one of the two patterns (*Parlando* or *Variable Foot*) is to create a text-speech alignment for the written poems and spoken recordings. In a second step, we detected the syntactic features, in particular the words' Part-of-Speech (PoS) in order to identify those poems using a "dismemberment of the line"[8] pp. 880 by separating the sentences into a nominal phrase and a verbal phrase. For the automatic identification of the rhythmical patterns mentioned above, we utilized the same tools already used in our method for enjambment detection in German readout poetry [9]:

- **Text-Speech Aligner**: The text-speech aligner [10], which implemented a variation of the SailAlign algorithm [11] using the Sphinx-4 speech recognizer [12], is utilized in this work.
- **Parser**: The Stanford parser [13] is used to parse the written text of poems. The main problem in poem parsing is the absence of punctuation in most cases, special characters as well as the writing in lowercase and the writing of other words such as articles in uppercase which cause errors into the parsing process.

2.3. Computational Analysis

The method employs an analysis based on computational speech processing in combination with manual philological analysis. The analysis is based on automatically extracted features to describe rhythmical patterns in *free verse poetry*. We parsed each line in the poem separately and expected that the number of lines including finite verbs - beside other possible non-finite verbs - is smaller than the total amount of lines. So whenever a full sentence containing one finite verb is spread over two or more lines, we have either a *Parlando* or a *Variable Foot*.

The Stanford parser uses for German the Stuttgart-Tübingen-TagSet (STTS) table [14] to identify the tags. We focused on the following verbs: finite verbs (VVFIN), imperative verbs (VVIMP), auxiliary verbs (VAFIN), auxiliary imperative verbs (VAIMP), and finite modal verbs (VMFIN). In a first step, finite verbs are located in each poetic line. Whenever a further finite verb was missing in the following line of the poem, then we had the "dismemberment of the line" [8] typical for both patterns to be detected: The *Parlando* as well as *Variable Foot*.

A further information for this identification of clauses below the sentence-level was taken from the existing punctuation marks. The complete sentences in lines can be identified by sentence ending punctuation (. ? ! ; :) and clauses by the comma. Therefore, all punctuation marks are detected in every poetic line. To differ between *Parlando* and *Variable Foot*, we compare the pause lengths between all the words in each individual line of the poem (P1) with the pause lengths between all the lines of the poem (P2) by using the text-speech alignment results. If P2 was greater than P1, then the poems rhythm is based on the *Variable Foot*.

Figure 1a shows an examplary analysis of the *Parlando*, identifying (from top to bottom): speech signal, intensity (dB), word alignment, end of line alignment, parser information (PoS-tagging), and time. This poem offers the syntactic features explained above: Most of the lines are based on a phrase/clause. In other words: The poem has more lines than finite verbs, not every line has a finite verb (identified by the PoS-Tagger as VVFIN or VAFIN). Figure 1a also shows that there is just a very short pause between the end of the first and the start of the second line (the pause length between the words "Gainsboroughs" and "wurde" is 0.4 sec). Compared to the analysis of the *Variable Foot* shown in Figure 1b below, the pause between the lines in Figure 1a is much shorter. This is obvious with regards to the time-layer: In Jandls poem coined by the *Variable Foot* the pause length between the two lines (between the words "lippen" and "ohne") is 1.1 sec.

2.4. Experimental Setup

For the automatic classification of *Parlando* vs. *Variable Foot*, different features including pause and parser information are extracted to build a model that differentiates these two classes. The following feature sets are utilized:

- **Pause**: the feature vector consists of two pause features: pause length between the words in each individual line and the pause length at the end of each line.
- **Parser**: three features from the parser are used: number of lines (lines with text), number of lines with finite verbs, and number of lines with punctuation.
- **Pause & Parser**: the current feature vector includes the five features used in the previous feature vectors: pause length between words in each individual line, pause length at the end of each line, number of lines, number of lines with finite verbs, and number of lines with punctuation.

A number of classifiers has been selected in order to determine the best suited classifier for the evaluation. The following machine learning algorithms with default values using the Weka data mining toolkit [15] are applied, whereas these classifiers are based on different classification techniques:

- **AdaBoostM1**: the boosting algorithm uses the Adaboost M1 method [16].
- **IBk**: the Instance-Based (IB) classifier with a number of (k) neighbors is the K-nearest neighbours (KNN) classifier using the euclidean distance and 1-nearest neighbour [17].
- **SimpleLogistic**: a classifier for building linear logistic regression models [18][19].
- **RandomTree**: Random trees is a collection of decision trees that considers K randomly chosen attributes at each node [20].

We used the above-mentioned classifier types and feature sets to identify those two rhythmical classes separated from the whole *lyrikline*-corpus. To deal with the low amounts of data, we use 10-fold cross-validation (each poem is present only in either the test- or the training folds).

3. Results

We identified the arithmetic average value for the syntactic features of poems by detecting the mean of the distribution in both

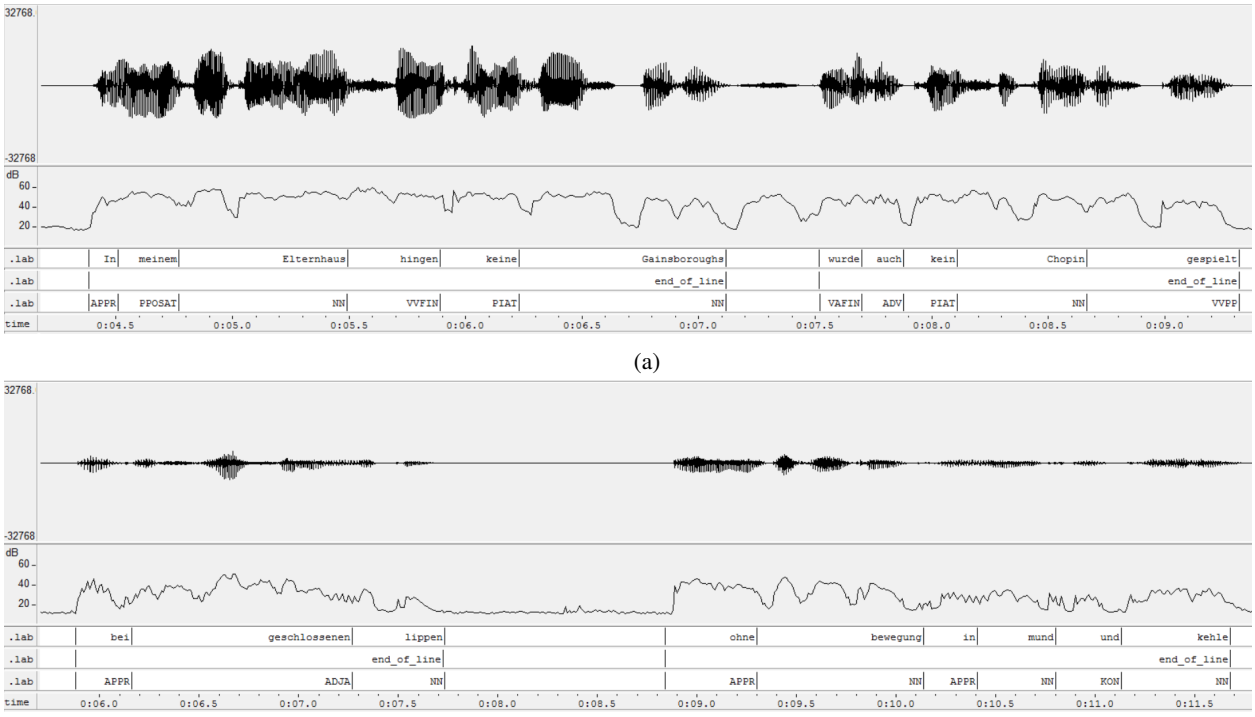

Figure 1: *(a) Analysis of the first two lines in the poem "TEILS-TEILS" (english: Half Here, Half There) from the poet "Gottfried Benn" as an example for the "Parlando" pattern.*
(b) As for (a) but for the poem "beschreibung eines gedichtes" (english: description of a poem) from the poet "Ernst Jandl" as an example for the "Variable Foot" pattern.

Table 1: *Experimental results (weighted average of f-measure) obtained with the 10 fold cross-validation by applying different feature sets on several classification algorithms.*

Classifier \ Features	Pause	Parser	Pause & Parser
AdaBoostM1	0.59	**0.69**	0.62
IBk	**0.68**	0.68	0.59
SimpleLogistic	0.47	0.63	**0.66**
RandomTree	0.65	0.56	0.53

classes. In the whole *Parlando*-corpus, we found per average 37 lines, 18 lines with finite verbs, and 25 lines using a punctuation. In the *Variable Foot*-corpus, the same distribution was 20, 10, and 11. This indicates that the poetic lines in both classes do hardly contain complete sentences and that these poems belong to both classes: *Parlando* and *Variable Foot*. The results of classifying poems as dominated by *Parlando* or *Variable Foot* are presented in Table 1. As can be seen, the performance is best with the AdaBoostM1 classifier using the parser information (f-measure is 0.69). There is a small difference in the best classification results using the different three feature vectors. The analysis of the decision tree using the classifier J48 [21] for the pause and parser features shows that the both features (number of lines and the number of lines with finite verbs) led to better cross-validation results in comparison to the other features.

4. Conclusion and Future Work

This paper is part of a research project trying to identify rhythmical patterns in modern and postmodern poetry by analyzing a huge corpus of readout poems collected from the *lyrikline* website. We compared these rhythmical features with tonality rules and focused on two important rhythmical patterns which use tonality-like features in poetic language (*Parlando* and *Variable Foot*). In the digital analysis, we used different features including pause and parser information to identify these two rhythmical patterns. In a further step, we extracted three feature vectors using pause, parser, and combination of pause and parser features. The classification with parser features yielded better results (f-measure is 0.69).

We already identified 17 different rhythmical patterns on *lyrikline* [22]. The main task of the project in the future will be to classify all of these rhythmical patterns by using the four features explained above. Therefore we will use a certain toolchain including PoS-tagger (for the grouping-structure), Tone and Break Indices ToBI (for the metrical structure), and Sonic Visualizer (for mapping time-span variations and prolongational structures). The last step will be the use of machine learning techniques to recognize all the rhythmical patterns in the *lyrikline* database.

5. Acknowledgements

The work is funded by the Volkswagen Foundation in the announcement 'Mixed Methods in the humanities? Funding possibilities for the combination and the interaction of qualitative hermeneutical and digital methods'.

6. References

[1] F. Lerdahl and R. Jackendoff, *A Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press, 1983.

[2] J. D. White, *The Analysis of Music*. New York: Prentice-Hall, Inc.,, 1976.

[3] E. Berry, "The free verse spectrum," *College English*, vol. 59, no. 8, pp. 873–897, 1997.

[4] E. Jandl, *der gelbe Hund. Gedichte*. Luchterhand Verlag, Darmstadt und Neuwied, 1982.

[5] E. M. Kafalenos, *Possibilities of Isochrony: A Study of Rhythm in Modern Poetry*. Washington, 1974.

[6] G. Benn, *Sämtliche Gedichte / Künstlerische Prosa - Band 1*. Klett-Cotta, Stuttgart, 2006.

[7] R. Cureton, *Rhythmic Phrasing in English Verse*. Longman, 1992.

[8] E. Berry, "The Free Verse Spectrum," *College English*, vol. 8, no. 59, p. 873–897, 1997.

[9] H. Hussein, B. Meyer-Sickendiek, and T. Baumann, "Automatic Detection of Enjambment in German Readout Poetry," in *Proceedings of the Speech Prosody*, Poznań, Poland, June 2018 (in press).

[10] T. Baumann, A. Köhn, and F. Hennig, "The Spoken Wikipedia Corpus Collection: Harvesting, Alignment and an Application to Hyperlistening," *Language Resources and Evaluation*, 2018.

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

[12] W. Walker, P. Lamere, P. Kwok, B. Raj, R. Singh, E. Gouvea, P. Wolf, and J. Woelfel, "Sphinx-4: A Flexible Open Source Framework for Speech Recognition," Mountain View, CA, USA, Tech. Rep., November 2004.

[13] A. N. Rafferty and C. D. Manning, "Parsing Three German Treebanks: Lexicalized and Unlexicalized Baselines," in *Proceedings of the Workshop on Parsing German*, ser. PaGe '08. Stroudsburg, PA, USA: Association for Computational Linguistics, 2008, pp. 40–46. [Online]. Available: http://dl.acm.org/citation.cfm?id=1621401.1621407

[14] A. Schiller, T. Teufel, C. Stöcker, and C. Thielen, "Guidelines für das Tagging deutscher Textcorpora mit STTS (Kleines und größes Tagset)," Available on http://www.ims.uni-stuttgart.de/forschung/ressourcen/lexika/TagSets/stts-1999.pdf, 1999, last accessed at 18. December 2017.

[15] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA Data Mining Software: An Update," *SIGKDD Explorations*, vol. 11, no. 1, pp. 10–18, 2009.

[16] Y. Freund and R. E. Schapire, "Experiments with a New Boosting Algorithm," in *Thirteenth International Conference on Machine Learning*. San Francisco: Morgan Kaufmann, 1996, pp. 148–156.

[17] D. Aha, D. Kibler, and M. Albert, "Instance-Based Learning Algorithms," *Machine Learning*, vol. 6, pp. 37–66, 1991.

[18] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," vol. 95, no. 1-2, pp. 161–205, 2005.

[19] M. Sumner, E. Frank, and M. Hall, "Speeding up logistic model tree induction," in *9th European Conference on Principles and Practice of Knowledge Discovery in Databases*. Springer, 2005, pp. 675–683.

[20] E. Frank and R. Kirkby, "Weka Classifiers Trees: Random Tree," Available on http://weka.sourceforge.net/doc.dev/weka/classifiers/trees/RandomTree.html, 2018, last accessed at 22. February 2018.

[21] R. Quinlan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers, 1993.

[22] B. Meyer-Sickendiek, H. Hussein, and T. Baumann, "Rhythmicalizer: Data Analysis for the Identification of Rhythmic Patterns in Readout Poetry (Work-in-Progress)," in *INFORMATIK 2017*, M. Eibl and M. Gaedke, Eds. Gesellschaft für Informatik, Bonn, 2017, pp. 2189–2200.