RECOGNIZING MODERN SOUND POETRY WITH LSTM NETWORKS

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

Abstract: Our paper focuses on the computational analysis of "readout poetry" (german: Hördichtung) – recordings of poets reading their own work – with regards to the most important type of this genre, the modern "sound poetry" (german: Lautdichtung). Whereas "readout poetry" often uses normal words and sentences, the "sound poetry", developed by dadaistic poets like Hugo Ball and Kurt Schwitters or concrete poets like Ernst Jandl, Oskar Pastior, or Bob Cobbing, combines the "microparticles of the human voice" like the segments in Ernst Jandls sound poem "schtzngrmm" ("schtzngrmm / schtzngrmm / tttt / tttt / grrrmmmmm / tttt / sch / tzngrmm"). Within the genre of sound poetry, there are two main forms: The lettristic and the syllabic decomposition. A short anecdote will explain this difference: The dadaist Raoul Hausmann developed the lettristic sound poetry in his early dadaistic poem "fmsbw" from 1918. This is said to have inspired his successor Schwitters, whose famous "Ursonate" [The Sonata in Primal Speech] begins with the words "Fümms bö wö tää zää Uu". With the "Ursonate", Schwitters developed a syllabic variation of the lettristic poems of Hausmann. The paper shows how to train a bidirectional LSTM network in order to differ between these "dadaistic" sound poems and the "normal" read out poems. In a further step, we will also show how to distinguish between the lettristic and the syllabic decomposition. Based on a bidirectional LSTM network that reads encodings of the character sequence in the poem and uses the output of each directional layer, we identify poems of the sound poetry genre and differentiate between its two types of decompositions. The classification results of sound poetry vs. other poetry as well as lettristic vs. syllabic decomposition are with a high performance, yielding a f-scores of 0.86 and 0.84, respectively.

1 Introduction

This contribution focuses on the computational analysis of "readout poetry" (german: Hördichtung) – recordings of poets reading their own work – with regards to the most important portal for international audio poems: the Berlin portal *lyrikline* (`www.lyrikline.org`). To date, readout poetry was never analyzed systematically, this is why there is no scientific knowledge about the prosodic features of this genre. We just know that at least 80 percent of the poems on *lyrikline* have neither a rhyme nor a fixed metre such as Jambs or Trochees. But does this mean that they are completely free of rhythmic structures? To answer this question, we make use of the so-called *free verse prosody*, a theoretical discourse developed in the US with regards to the modern poetry developed by poets such as Whitman [1][2], the Imaginists [3][4][5], the Beat poets or today's Slam poets. All these authors have replaced the classical metric verse forms with a new prosody that is characterized by prose rhythms, everyday language, linguistic decompositions like in modern "sound poetry" (german: Lautdichtung) or musical styles such as Jazz or Hip-Hop.

To date, we found 17 different rhythmical patterns [6], to be divided into 4 different categories: textual patterns, acoustic patterns, experimental patterns and decompositions. In this paper we will focus on those patterns to be found in sound poetry (decompositions). Generally speaking, sound poetry intends to limit the poetic language to the “microparticles of the human voice”, as Bob Cobbing once put it [7]. Skipping the syntactic as well as semantic relations, sound poetry regularly excludes the level of meaning by reducing the texts to the arrangement of letters, sequences of letters and syllables according to specific patterns or series. This technique can be varied with regards to the extension of the alphabetical inventory by new characters or letters, a technique used by authors such as Isidore Isou, Maurice Lemaître, or Valeri Scherstjanoi. A further technique in sound poetry is the anagram or the palindrome, often used by authors like Bob Cobbing, Josef Anton Riedl or Gerhard Rühm. The most important example for these techniques are the poems of Oskar Pastior, who also emphasizes the nature of language: he is not only concerned with the semantics of language, but also with the visual and acoustic form of words or verses. Pastior uses in his work techniques such as anagrams, palindromes or lipograms, in the volume “Kopfnuß, Januskopf” from 1990, he compiles texts based on the palindrome, i.e. a string identical to read both forwards and backwards.

Using electronic techniques, the french poet Henri Chopins created his Poésie sonore, further examples are text sound compositions of the Swedes Lars-Gunnar Bodin and Bengt Emil Johnson as well as Carlfriedrich Claus in his sound units. On the portal *lyrikline*, there are sound poems from both kinds, created by authors like the Australian Amanda Steward, the Dutchman Jaap Blonk and the Russian Valeri Scherstjanoi, probably the best-known artists in this field, and Michael Lentz, himself author and interpreter of sound poetry, and publisher of a comprehensive inventory [8][9].

In this paper, in order to classify sound poetry, we isolated at first this type of poetry – sound poetry – from the other ‘normal’ poetry. Then, we classified both forms of sound poetry (lettristic and syllabic decomposition). For this purpose, we trained a model of the character sequences by using a bidirectional Long Short-Term Memory (LSTM) network.

The paper is organized as follows: Section 2 gives an overview on the project *Rhythmicalizer*. The database is presented in the Section 3. Section 4 reviews the philological and technical method for the automatic recognition of sound poetry. The experiment and results are described in Section 6. Finally, conclusions and future works are presented in Section 7.

2 Rhythmicalizer

The aim of the project *Rhythmicalizer* (`www.rhythmicalizer.net`) is the automatic recognition of rhythmical patterns in modern and postmodern poetry. The used database is from our partner *lyrikline*. The database contains speech and text data of modern and contemporary poetry, giving us access to hundreds of hours of author-spoken poetry. The three-year project is funded by the Volkswagen Foundation (german: VolkswagenStiftung) in the ‘Mixed Methods’ program in the humanities. We want to develop a software for the digital classification of prosodic patterns by using the acoustic and textual data in readout poetry.

3 Database

The data used in the project is a large collection of modern and postmodern readout poetry taken from our partner *lyrikline*. The *lyrikline* hosts contemporary international poetry as audio files (read by the authors themselves) and texts (original versions & translations). Users can listen to the poet and read the poems both in their original languages and various translations. The digital material covers more than $10,800$ poems from more than $1,200$ international poets from more than 70 different languages. Nearly 80% of the *lyrikline*-poems are postmetrical poems. In our project, we will use all the poems in English and German languages (more than 3,600 poems). The total number of German and English poets is 215 and 154, respectively.

In this study, the philological scholar of our project collected only the German poems from the *lyrikline*-website and from different volumes of famous sound poets already mentioned above. The poems of sound poetry from a total of 11 poets are used in this study (6 poets by lettristic decomposition and 10 by syllabic decomposition). This selection is about 5% of all the German speaking poets. Each of these authors read at least one and at most 8 poems by both lettristic and syllabic decompositions, so we analyzed a total of 39 poems of sound poetry (16 by lettristic decomposition and 23 by syllabic decomposition). In addition, the philological scholar collected the "normal" German read out poems from the *lyrikline*-website from a total of 46 poets. This selection is about 21% of all the German speaking poets. Each of these authors read at least one and at most 9 poems, so we analyzed a total of 111 poems of "normal" readout poetry. We used only the text data in this experiment.

4 Method

Our Method is based on the interplay of hermeneutical and computational approaches. For this reason, we made use of a statistical method developed to analyze the modern *free verse poetry*. This method evaluates the relation between grammatical and metrical units in poetry. The automatic recognition of sound poetry is based on the training of a bidirectional LSTM network which uses the sequence of characters in the poem.

4.1 Lettristic and Syllabic Decompositions in Modern Sound Poetry

To classify the patterns of sound poetry systematically, we make use of the theory of the grammetrical ranking. The term grammetrics, coined by Donald Wesling, is a hybridization of grammar and metrics: The key hypothesis is that the interplay of sentence-structure and line-structure can be accounted for more economically by simultaneous than by successive analysis [10]. In poetry as a kind of versified language, the singular sentence interacts with verse periods (syllable, foot, part-line, line, rhymed pair or stanza, whole poem), a process for which Wesling finds 'scissoring' an apt metaphor: Grammetrics assumes that meter and grammar can be scissored by each other, that the cutting places can be graphed with some precision. One blade of the shears is meter, the other grammar. When they work against each other, they divide the poem. It is their purpose and necessity to work against each other [10].

In Wesling's scheme (see Figure 1), the vertical axis designates the grammatical and the horizontal axis the metric rank. Intersections of the two axes are represented by circles in which the axes meet; small circles for small coordinate points, large circles for large ones. Of all possible intersections on the grid, only 16 points are encircled in a black colour, because these 16 points are filled in "normal" readout poems. The two clusters illustrating this focus on the main coordinates are word and foot as well as sentence and line (large black circles).

The features of sound poetry - encircled in red - are localized on the smaller intersections on the grid, marking the units below the word level (< 3) on the grammatical rank, and below the foot level (< 3) on the metrical rank. Sound poetry can thus be regarded as a detachment from the language in which phonemes or letters move into the center of the poem. With regards to the grammetrical ranking, the poems participating in the sound poetry could use syntactic features, but they will try to reduce these features to the phonemic or syllabic level. The matrix shows that sound poetry differs from traditional poetry, because in traditional poetry "the word dominates" [11], as the famous Austrian poet Ernst Jandl puts it. The first lines of the famous

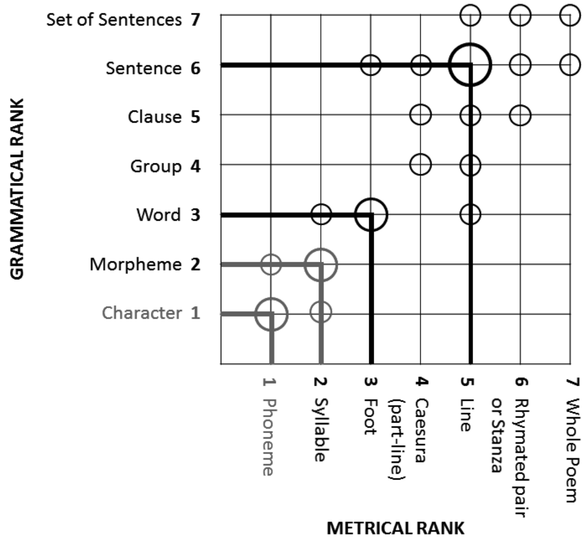

Figure 1 – The vertical axis is that of grammatical rank; the horizontal axis is that of metrical rank. Intersection points help to identify the poem, for instance, the line arrangement. The red color is an extension of Donald Wesling's original scheme in order to include the sound poetry features (after [10])

.

"Ursonate" from Kurt Schwitters – "Fümms bö wö tää zää Üü" – would be located on the first big red circle (grammatical rank = 2; metrical rank = 2), representing a typical example for the syllabic decomposition, whereas the lettristic decomposition in Raoul Hausmann's dadaistic poem, "fmsbw" could be located on the big red circle below (grammatical rank = 1; metrical rank = 1) [12]. A further example can be found in Ernst Jandls sound poetry. In his famous volume "Laut und Luise" [13], Jandl expanded the dadaist sound poetry by using both forms of decomposition (lettristic and syllabic). On the one hand, he devided the poetic words into syllables, imitating dadaistic sound poems of Hugo Ball and Kurt Schwitters. He also developed the lettristic kind of sound poetry by using a prosodic repetition of single characters like Raouls Hausmann in "fmsbw" on the other hand. Giving an example of both types of poetic decomposition in his lecture "Voraussetzungen, Beispiele und Ziele einer poetischen Arbeitsweise" [14], Jandl mentions his poems "schtzngrmm" and "auf dem land":

schtzngrmm (lettristic decomposition)
schtzngrmm
schtzngrmm
t-t-t-t
t-t-t-t
grrrmmmmm
t-t-t-t
s———c———h
tzngrmm
tzngrmm
tzngrmm
grrrmmmmm
schtzn
schtzn
t-t-t-t

auf dem land (syllabic decomposition)
rininininininininDER
brüllüllüllüllüllüllüllüllEN
schweineineineineineineineinE
grunununununununununZEN
hunununununununununDE
bellellellellellellellellEN
katatatatatatatatZEN
miauiauiauiauiauiauiauiauEN
katatatatatatatatER
schnurrurrurrurrurrurrurrurrEN
gänänänänänänänänSE
schnattattattattattattattattERN
ziegiegiegiegiegiegiegiegEN
meckeckeckeckeckeckeckeckERN

Both poems belong to the sound poetry, although they differ with regards to the grammetrical ranking: Whereas in “schtzngrmm” each line consists only of consonants, the poem “auf dem land” uses syllables in each line. Both these types of decomposition are used in sound poetry: Whereas dadaistic poets as well as the Vienna Group [15] are more focused on syllabic decomposition, the lettristic decomposition is typical for the works of Isidore Isou and Maurice Lemaître, whose poetry is based on individual letter sounds [16].

5 Character Sequences as Indicators for Sound Poetry

Each poem can be understood as a sequence of words, syllables or letters. As outlined above, sound poems hardly use words, but are based on a certain combination of syllables or characters, and their respective sounds. To detect these features automatically, different techniques could be taken into account. We could, in principle, perform look-ups against dictionaries and declare a poem sound poetry if these look-ups fail frequently. We could also compute the cross-entropy of a poem against a character N-gram model of ‘normal’ written German and would expect higher cross-entropy (i.e., larger deviations) for sound poetry than other poetry. We instead opt to distinguish the types of sound poetry mentioned above from other poetry based on regular sentences, by training a model of the character sequences found in sound poetry vs. other poetry. We train a bidirectional LSTM network that reads encodings of the character sequence in the poem and uses the output of each directional layer to predict whether the poem is sound poetry based on a final decision layer. When comparing our neural learning approach over a method based on dictionary look-ups or cross-entropy compared to standard German character models, our method is more general as it does not make use of external data, and can be retrained for other tasks. In fact, we can use the same model architecture to train a model that differentiates lettristic and syllabic sound poetry, a task on which the dictionary look-up would fail completely.

Given our very small database and the risk of overfitting, we cut our poems into line-by-line segments that are used for training and evaluation separately. We then declare a poem correctly classified if more than half of the lines in the poem are correctly classified. In addition, our line-based model allows us to perform fine-grained analyses of the poems, e.g. we can automatically identify parts of a lettristic poem that could also be characterized as syllabic.

6 Experiment and Results

Our neural network architecture is shown in Figure 2. Characters are encoded into 20-dimensional character embeddings which force to learn a coarse-grained differentiation of letters and helps to avoid overfitting to lexical characteristics. We use bidirectional two-layer LSTMs to encode each line of each poem and use the output states of each LSTM as input to a final decisioning layer which is followed by the two-dimensional softmax which determines the class output.

Given the very limited amount of training data, we use leave-one-out testing for every poem in the dataset (i. e. rather than splitting out a fixed testset, we consider every individual poem as testset and train the classifier based on all other poems, for every poem). For testing, we recombine all the line-by-line classifications and choose the class that is assigned to the majority of lines as the poem's class. Our system is implemented in *dyNet* [17] and we stop training after 10 epochs (iterations through the training data).

We performed two experiments:

- **Sound poetry vs. other poetry**: First we learn how to differentiate sound poetry from other “normal” poetry. For this, we selected poems previously collected from *lyrikline*,

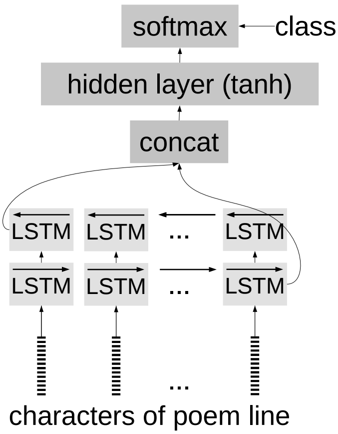

Figure 2 – Architecture of the neural network used for identifying poetry types. We use two-layer LSTMs with 10-dimensional hidden states, 20-dimensional character embeddings, and a 20-dimensional decision layer.

Table 1 – Classification results for the character sequence model.

	lines correct	poems correct	precision	recall	f-score
sound poetry identification	81 %	87 % (66/76)	.97	.77	.86
lettristic/syllabic differentiation	74 %	79 % (31/39)	.86	.83	.84

that differ from sound poetry, cause they use regular words or sentences, sometimes structured by elliptic phrases or line-breaks resp. enjambments. There is a total of 111 fully processed poems of this more or less “normal” shape and we randomly sampled a third, yielding 37 poems (as compared to the 39 within the sound poetry category).

- **Lettristic vs. syllabic decomposition**: As a second experiment, we learn how to differentiate the two classes of sound poetry: dominated by lettristic or syllabic decomposition, as outlined in the previous section. There are 16 poems dominated by lettristic decomposition and 23 by syllabic decomposition.

The results of the experiment are shown in Table 1. We find that both classification tasks (sound poetry vs. other poetry as well as lettristic vs. syllabic decomposition) are learned with reasonably high performance, yielding a f-scores of 0.86 and 0.84, respectively, and performing highly significantly above chance level (binomial test, $p \ll .01$).

We also find that our ad-hoc ‘the winner takes all’-approach to classify poems by the majority vote on lines performs robustly and is able to boost full-poem accuracy beyond the line-by-line accuracy, while greatly improving training performance (as evidenced in initial trials of the experiments).

In a final step, the philological expert looked at the results and analysed the erroneous output of the system. Many false results could not be explained easily and seem to occur randomly. For example, while “schtzngrmm” was correctly classified as lettristic decomposition, “auf dem land” was not, indicating that our model is still lacking generalization. In other cases, errors can be attributed to the decomposition being only observable in parts of the poem and only those lines are identified correctly, but do not make up the majority of lines in the poem.

7 Conclusion and Future Work

This paper presented an automatic method for the recognition of sound poetry within a large corpus of modern poetry, and for the classification of the most important features of sound poetry (the lettristic and syllabic decomposition). These two features are distinguished by their different combination of linguistic microparticles: on the one hand syllables, on the other hand letters. In this study, we focused on the textual patterns of sound poetry. We trained a bidirectional LSTM network to read encodings of the character sequence from every line in the poem. The classification results (f-score) for sound poetry vs. other poetry was 0.86; for lettristic vs. syllabic decomposition it was 0.84.

We assume that we get even better results when we make use of the speech signal as well. Our current alignment system does not work well for sound poetry which is why we leave out acoustic analysis in this work. We could, however, manually segment the poem's audio into lines and make use of the per-line audio material as additional support in our present model. We expect that Mel-Frequency Cepstral Coefficients (MFCCs) features will help us to get better results. As long as sound poems use either lettristic or syllabic decompositions, we expect to recognize this difference by focussing on the spectral features. There are far more vowels in the syllabic than in the lettristic sound poetry, this is why MFCC features will help us to improve our results. In addition, LSTMs should well be able to recover the recurrent structure of syllabic decomposition.

8 Acknowledgements

This 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' (funding code 91926 and 93255).

References

[1] GATES, R. L.: *The Identity of American Free Verse: The Prosodic Study of Whitman's 'Lilacs'*. *Language and Style*, (18), pp. 248—-276, 1985.

[2] GATES, R. L.: *T. S. Eliot's Prosody and the Free Verse Tradition: Restricting Whitman's "Free Growth of Metrical Laws"*. *Poetics Today*, 11(3), pp. 547–578, 1990.

[3] SILKIN, J.: *The Life of Metrical and Free Verse in Twentieth-Century Poetry*. Palgrave Macmillan UK, 1997.

[4] COOPER, G. B.: *Mysterious Music: Rhythm and Free Verse*. Stanford University Press, 1998.

[5] BEYERS, C.: *A History of Free Verse*. University of Arkansas Press, 2001.

[6] MEYER-SICKENDIEK, B., H. HUSSEIN, and T. BAUMANN: *Rhythmicalizer: Data Analysis for the Identification of Rhythmic Patterns in Readout Poetry (Work-in-Progress)*. In M. EIBL and M. GAEDKE (eds.), *INFORMATIK 2017*, pp. 2189–2200. Gesellschaft für Informatik, Bonn, 2017.

[7] COBBING, B.: *Poetry for a new age*. No. 2, p. 34. Edited by Dom Sylvester Houédard, London, 1971.

[8] LENTZ, M.: *Laut-Poesie/-musik: nach 1945*. Laut-Poesie/-musik nach 1945: eine kritisch-dokumentarische Bestandsaufnahme. Ed. Selene, 2000.

[9] LENTZ, M.: *MUSIK? POESIE? EIGENTLICH...: LAUT POESIE MUSIK NACH 1945*. *Neue Zeitschrift für Musik (1996)*, 157(2), pp. 47–55, 1996.

[10] WESLING, D.: *The Scissors of Meter: Grammetrics and Reading*. University of Michigan Press, 1996.

[11] JANDL, E.: *Autor in Gesellschaft: Aufsätze und Reden*. p. 217. Luchterhand Verlag, Darmstadt, 1999.

[12] PETERS, M.: *Die mißlungene Rettung der Sonate. Kurt Schwitters' Ursonate als neue Vokalmusik*. pp. 217–223. MUSICA, Kassel, 1977.

[13] JANDL, E.: *Laut und Luise. Mit einem Nachwort von Helmut Heißenbüttel*. Luchterhand Verlag, Darmstadt, 1971.

[14] JANDL, E.: *Voraussetzungen, Beispiele und Ziele einer poetischen Arbeitsweise*. pp. 25–40. Wien-München, 1970.

[15] KONIECZNA, E.: *Die deutsche Dichtung heute (Eine Darstellung unter besonderer Berücksichtigung der konkreten Poesie in Österreich)*. No. 2, pp. 29–41. Filologia Germanska, Tornn, 1976.

[16] CURTAY, J.-P.: *La poésie lettriste*. Seghers, Paris, 1974.

[17] NEUBIG, G., C. DYER, Y. GOLDBERG, A. MATTHEWS, W. AMMAR, A. ANASTASOPOULOS, M. BALLESTEROS, D. CHIANG, D. CLOTHIAUX, T. COHN, K. DUH, M. FARUQUI, C. GAN, D. GARRETTE, Y. JI, L. KONG, A. KUNCORO, G. KUMAR, C. MALAVIYA, P. MICHEL, Y. ODA, M. RICHARDSON, N. SAPHRA, S. SWAYAMDIPTA, and P. YIN: *Dynet: The dynamic neural network toolkit*. *arXiv preprint arXiv:1701.03980*, 2017.