



# Exploring intonational patterns of poetic speech: Insights from a large corpus of German poetry

Nadja Schauffler<sup>1</sup>, Nora Ketschik<sup>1</sup>, André Blessing<sup>1</sup>, Kerstin Jung<sup>1</sup>, Toni Bernhart<sup>1</sup>, Anna Kinder<sup>2</sup>, Julia Koch<sup>1</sup>, Jonas Kuhn<sup>1</sup>, Sandra Richter<sup>1,2</sup>, Rebecca Sturm<sup>2</sup>, Gabriel Viehhauser<sup>1</sup>, Thang Vu<sup>1</sup>

<sup>1</sup>University of Stuttgart, Germany

<sup>2</sup>German Literature Archive Marbach, Germany

nadja.schauffler@ims.uni-stuttgart.de

## Abstract

Poetic speech has often been described to differ from readings of prose and other texts in a number of features. While so far, investigations of poetic speech have mostly covered only small samples of speakers and poems, the study at hand investigates prosodic features in a large corpus of over 1400 German poetry recitations spanning from the 1950s to the 2000s. In particular, we investigate how factors such as the length of the poem, authorship, and time of recitation influence speech rate and pitch variation. Our results show significant changes in recitation rate and pitch range over time, indicating diachronic trends in recitation patterns that can also be found within individual speakers. Additionally, we observe that the poem's author affects how the poem is performed, with some authors being recited at faster rates and with greater pitch variation than others. The findings will serve as the basis for a refined poetic speech synthesis model capable of capturing stylistic nuances related to speakers, authors, or recitation trends, allowing a deeper exploration of their aesthetic effects.

**Index Terms:** poetry, speech style, diachronic corpus analysis, speech rate, pitch variation

## 1. Introduction

This paper presents findings on prosodic features of poetic speech from a large collection of German recitations. The fact that poetic speech differs from other kinds of speech (spontaneous speech, prose reading, broadcast) has been addressed in multiple studies across languages and disciplines. Byers [1] compared poetry readings to prose and conversation samples by the same six speakers. She identified a number of intonational features typical of verse recitation in English and modeled them in what she called "a formula for poetic intonation". This "formula" comprised features such as a slow speech rate, short intonation units, more pauses, intonation units of relatively equal length, low average pitch, a narrow pitch range, simple falling melodies, and simple falling nuclear tones. Comparing one poetry recording with one radio news broadcast, Barney [2] identified a "new formula for poetic intonation" that categorized short tone units, more pauses, units of relatively equal length and echoes between pitch patterns as specific poetic features, while slow speech rate, narrow pitch range, and low average pitch were considered general performance features.

The role of pausing, lengthening and speech rate as poetic speech features has recently been investigated in other languages such as German [3, 4] and Portuguese [5]. For German, research has shown that poem recordings are slower compared to prose, due to the longer duration of speech units and the more frequent occurrence of silent pauses [3, 4]. Additional findings from eye-tracking confirmed these genre-dependent dif-

ferences, revealing that readers read more slowly in poetry and are less inclined to skip words than in prose [3]. With the aim to provide speech synthesis models with a fixed metre, Bröggelwirth [4] analyzed the rhythm of poetic speech and confirmed that listeners were able to differentiate between the speaking style of poetry and prose based on reduced data (delexicalized, monotonized).

Wagner and Betz [6] further investigated how the speaker's experience influenced aspects of prosodic timing showing that experienced speakers systematically used pauses and prosodic boundaries while laypersons did not rely on the same time management measures.

In summary, previous studies primarily compared features of poetic speech to non-poetic speech, such as prose or broadcast, often using small sample sizes. Some studies also explored connections between specific poetic speech features and speaker characteristics, such as experience. Advances in computational methods now allow for the analysis of larger corpora, offering new opportunities for research in this area. The study at hand is based on a large and richly annotated poetry corpus which enables us to explore prosodic aspects of poetic speech across different authors, speakers and temporal influences. Specifically, we analyze how pitch variation, speech rate and pausing are affected by the length of a poem, authorship or speaker and we gain a diachronic perspective on the data by looking at trends in the recitation of poetry over time.

## 2. Material and Methods

### 2.1. Corpus

The corpus for the study at hand is part of the data collected within the project »textklang« and comprises 1408 recordings of 1005 German poems by 54 authors, spoken by 117 different speakers (36 female) between 1902 and 2020. Most of the poems were written between approximately 1745 and 1900, with a focus on the Romantic period. Therefore, the corpus constitutes a relatively homogeneous collection with respect to the epoch. This allows us to identify speaker dependent differences, stylistic trends in declaration as well as particularities of specific authors without taking into account crucial epochal influences.

### 2.2. Annotations

The corpus was manually enriched with metadata such as the speaker, time and context of recitation, time of origin and author of the poem. Automatic annotations for text and audio

<sup>1</sup>The interdisciplinary project »textklang« develops a mixed-methods approach for the systematic investigation of the relationship between written text and its sonic realization, cf. <http://hdl.handle.net/11022/1007-0000-0007-F6C5-5>

largely followed the pipeline as developed for the GRAIN corpus [7] with additional annotation layers for poem-specific features such as lines and stanzas (cf. also [8] for a list of all metadata and automatic annotations). The acoustic data was force aligned for phone, syllable, and word boundaries [9]. Pauses, pitch and duration values for each syllable were extracted by means of the Festival [10] version of the University of Stuttgart [11], which is a synthesis system for German.

### 2.3. Analyses

We investigated how the length of the poem, the time of recitation, and the author of the poem influenced the speech rate and variation in pitch of a poem’s recitation. Given that there were only a few recitations from some periods of time, we only kept decades with more than 30 recitations for the statistical analyses. This excluded recitations from the 1940s and earlier, as well as recitations recorded after 2000 and left 1380 recordings to analyze.

We performed linear mixed-effects analyses to predict the following features:

- The frequency of pauses, calculated as the number of silences divided by the number of lines.
- The articulation rate, calculated as the number of syllables per second (excluding pauses).
- The recitation rate, calculated as the number of syllables per second over the length of the respective recitation (this includes pauses within but also across stanzas).
- Pitch range, referring to the pitch span calculated as the difference between the lowest and highest pitch within one line, taken from the middle of the respective syllable.<sup>2</sup>

Paratexts, such as lines with the title or the author’s name, were excluded from all analyses. Pitch values below the 1st and above the 99th percentile per speaker were excluded as outliers before pitch range was calculated. This removed 1.37 % of the data and left 38,692 lines to analyze. We calculated the mean value of the line-by-line pitch ranges per recitation, so that there was one value (in Hz) per item. As random intercepts, we used speaker and poem. As fixed factors, we used the following:

- The length of the poem as the number of lines. Poem length ranges from 2 lines to 257 lines. In order to avoid a disproportionate influence of very long poems of which there are only few, the line count was log-transformed and afterwards centered by subtracting its mean from each value.<sup>3</sup>
- The time of recitation categorized in decades.

To investigate the effect of authorship, we used a subset of the data comprising recitations of the six most frequently recited authors ( $n > 70$  recitations). This subcorpus comprised 995 recordings of poems by Joseph von Eichendorff, Johann Wolfgang von Goethe, Heinrich Heine, Friedrich Hölderlin, Eduard Mörike, and Friedrich Schiller. We performed linear mixed models on this subcorpus again with pauses, recitation rate and pitch range as dependent variables, respectively, author and line count as fixed factors and only speaker as random intercept (since poem and author naturally correlate).

All statistical analyses were performed in R 4.2.2 [12], using the *lmer* function of the packages *lme4* and *lmerTest*

<sup>2</sup>We use the pitch range over the line as a means to represent how “expressive” and modulated or varied the recitation is, since the line is the organizing unit in poems. This means that pitch differences may be measured across pauses and do not represent pitch spans of prosodic phrases.

<sup>3</sup>One poem with 300 lines was excluded as outlier.

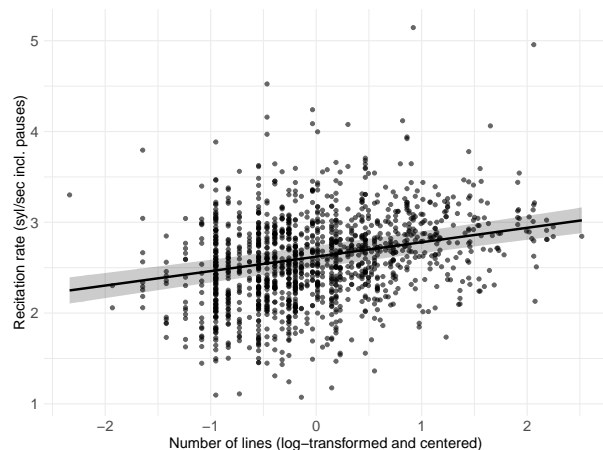


Figure 1: *The effect of poem length of poem with the log-transformed and centered line count on the x-axis and the recitation rate (syl/sec incl. pauses) on the y-axis.*

[13, 14]. All factors were tested for significance using likelihood ratio tests by comparing the model that includes the effect in question with the model that does not. Post-hoc comparisons were performed by calculating contrasts using the *emmeans* package (Tukey method) [15], plots were generated using *effect\_plot* of the package *jtools* [16] and *ggplot2* [17].

## 3. Results

Due to space limitations, we only report and discuss effects for recitation rate (as a combination of pausing and articulation rate) and pitch range. All results and corresponding data (without the licenced recordings) can be made available upon request.

### 3.1. Recitation rate

First, we analyzed the influence of poem length, year of recitation, and author of poem on the recitation rate. Figure 1 shows the effect of the poem’s length as predicted by the model: The longer the poem, the faster it is recited ( $\beta = 0.16$ ,  $SE = 0.02$ ,  $t = 9.58$ ,  $p < .0001$ ).

Additionally, we found an effect of decade: Recitations from the 1980s were recited more slowly than earlier recitations from the 1960s ( $\beta = 0.19$ ,  $SE = 0.05$ ,  $t = 3.89$ ,  $p < .001$ ) and later recitations from the 2000s ( $\beta = 0.20$ ,  $SE = 0.07$ ,  $t = 3.08$ ,  $p < .01$ ), as can be seen in Figure 2. Note, however, that while there were a number of recitations from the 1980s (125), these were by only three speakers, pointing towards a possible speaker effect. We are going to come back to that issue in 3.3.

Examining a subcorpus comprising the six most frequent authors in the corpus ( $n > 70$  recitations), we also observed that recitation rate is affected by the poem’s author. Two authors in particular are read at a significantly higher rate, as can be seen in Figure 3: Heine compared to Eichendorff ( $\beta = 0.25$ ,  $SE = 0.06$ ,  $t = 3.98$ ,  $p < .01$ ), to Goethe ( $\beta = 0.21$ ,  $SE = 0.06$ ,  $t = 3.40$ ,  $p < .01$ ), to Hölderlin ( $\beta = 0.19$ ,  $SE = 0.07$ ,  $t = 4.08$ ,  $p < .0001$ ) and to Mörike ( $\beta = 0.19$ ,  $SE = 0.06$ ,  $t = 3.03$ ,  $p < .0001$ ); and Schiller compared to Eichendorff ( $\beta = 0.23$ ,  $SE = 0.07$ ,  $t = 3.26$ ,  $p < .05$ ), to Goethe ( $\beta = 0.19$ ,  $SE = 0.07$ ,  $t = 2.8$ ,  $p < .05$ ), and to Hölderlin ( $\beta = 0.30$ ,  $SE = 0.07$ ,  $t = 4.04$ ,  $p < .001$ ).

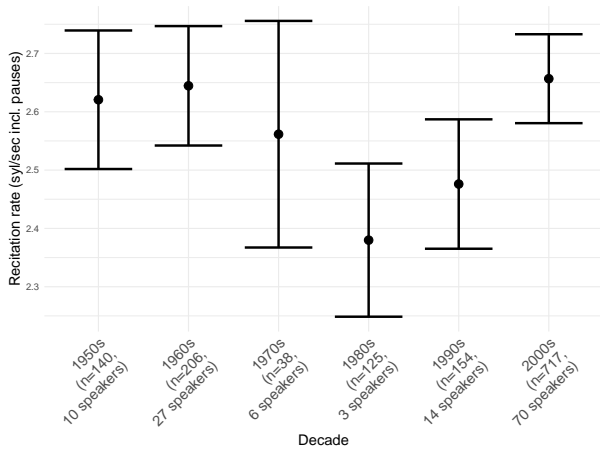


Figure 2: The effect of time of recitation on the x-axis with number of recitations and speakers per decade and the recitation rate (syll/sec incl. pauses) on the y-axis.

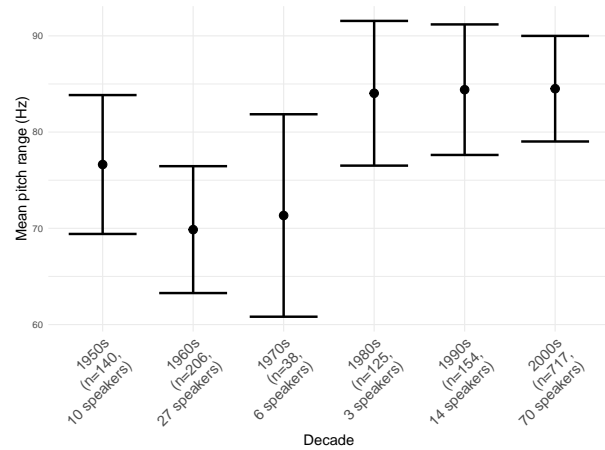


Figure 4: The effect of decade on the x-axis with number of lines and speakers, and the mean pitch range on the y-axis.

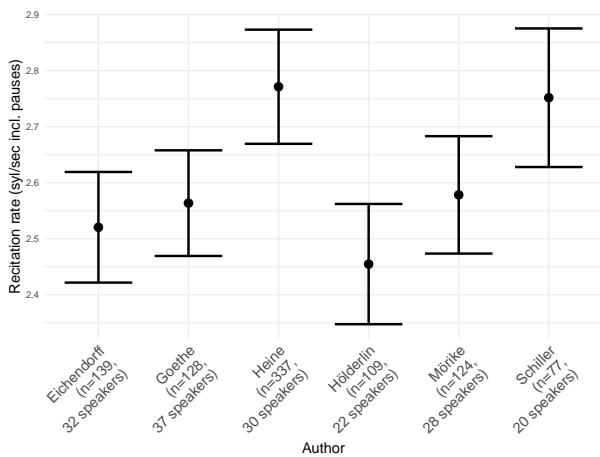


Figure 3: The effect of author on the x-axis with number of recitations and speakers, and the recitation rate (syll/sec incl. pauses) on the y-axis.

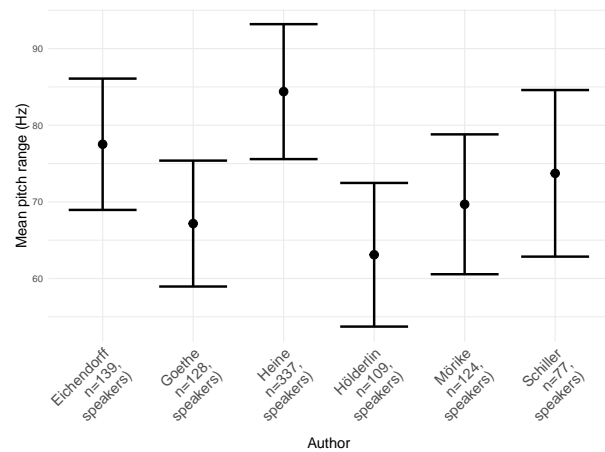


Figure 5: The effect of author on the x-axis with number of recitations and speakers, and the mean pitch range on the y-axis.

### 3.2. Pitch range

In order to investigate how the length of the poem and the authorship as well as the time of recitation affect how "expressive" or varied and modulated the speaker performs the text, we predicted the mean pitch range by decade and poem length.

Figure 4 shows the effect of decade on the pitch range as predicted by the model. Speakers of earlier recitations recited poems with a smaller pitch range compared to speakers of later recitations. Recitations of the 1950s had a significantly smaller mean pitch range compared to recitation in the 1990s ( $\beta = -7.76$ ,  $SE = 2.57$ ,  $t = -3.02$ ,  $p < .05$ ). And recitations of the 1960s had a significantly smaller range compared to the 1980s ( $\beta = -14.17$ ,  $SE = 2.44$ ,  $t = -5.80$ ,  $p < .0001$ ), the 1990s ( $\beta = -14.45$ ,  $SE = 2.32$ ,  $t = -6.26$ ,  $p < .0001$ ) and the 2000s ( $\beta = -14.64$ ,  $SE = 3.42$ ,  $t = -4.27$ ,  $p < .001$ ). There was no effect of the poem's length on the mean pitch range.

Looking at the subset comprising the most frequently recited authors in the corpus, we observed that the pitch span partly differed depending on the poem's author, cf. Figure 5.

We found that Heine was recited with a significantly wider mean pitch range compared to poems by Goethe ( $\beta = 17.22$ ,  $SE = 5.44$ ,  $t = 3.17$ ,  $p < .001$ ), and compared to poems by Hölderlin ( $\beta = 21.22$ ,  $SE = 5.84$ ,  $t = 3.63$ ,  $p < .01$ ).

### 3.3. Change of style in one speaker

Considering the change in recitation rate found for the 1980s, we come back to the issue of having only three speakers for that time period. Out of 125 recitations in the 1980s, 93 are from a single speaker, Gert Westphal. This speaker, an actor, recited poems from the 1960s to the 2000s. This is why we created another subset, comprising only Westphal's recitations, in order to find out whether the effect of decade was mainly caused by this speaker or whether it represents a particular trend during that time. We categorized Westphal's recitations into "early", comprising recordings from 1961 to 1965 ( $n = 49$ ), and "late", with recordings from 1986 to 1992 ( $n = 131$ ). We performed a linear mixed model with recitation rate as dependent variable, time of recitation (early vs. late) as fixed factor and poem as a random intercept. We found that he performed his later recita-

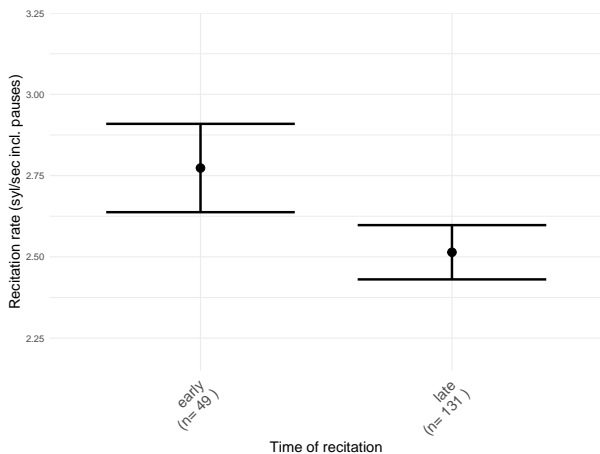


Figure 6: *The effect of early (1960s) vs. late (1980s/1990s) on the x-axis, and the recitation rate (syl/sec incl. pauses) on the y-axis for the speaker Gert Westphal.*

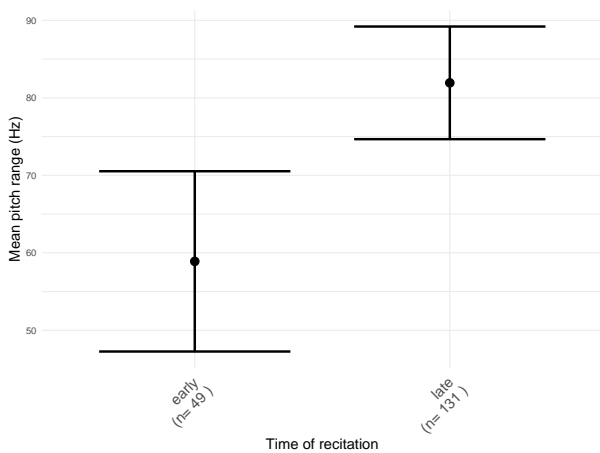


Figure 7: *The effect of early (1960s) vs. late (1980s/1990s) on the x-axis, and the mean of pitch range (mean of line-wise pitch range in Hz) on the y-axis for the speaker Gert Westphal.*

tions significantly more slowly than the earlier ones ( $\beta = -0.26$ ,  $SE = 0.08$ ,  $t = -3.27$ ,  $p < .01$ ), as can be seen in Figure 6. In addition, we observed a change in his use of pitch variation: In his later recitations, the mean pitch range is significantly wider than in his earlier recitations ( $\beta = 23.04$ ,  $SE = 6.68$ ,  $t = 3.46$ ,  $p < .001$ ), see Figure 7. Both results reflect the pattern found for the entire corpus.

#### 4. Discussion and Conclusion

We analyzed a large corpus of German poetry recitations to examine how prosodic features typical of poetic speech are influenced by factors such as poem length, authorship, and the time of recitation. While there is considerable variation within the investigated features, we identified speaker-independent factors affecting the performances. In terms of formal characteristics, our findings indicated that longer poems tended to be recited at a higher rate. Comparing a subset of the six most frequently recited authors in the corpus, we noticed that it also played a role by whom the poem was written. Poems by Heine and Schiller,

in particular, were read at a significantly higher recitation rate. Heine was also found to have a significantly wider mean pitch range compared to other authors. This is interesting insofar as the stylistic characteristics of the authors seem to be deliberately conveyed by the recitation, regardless of who recites the poem and at what point in time. Heine, for example, is characterized by a certain (ironic and prosaic) tone which seems to be transferred into the recitation of his poems. Also Schiller's poems – namely his ballades – are characterized by narrative elements (since ballades are considered "narrative poems"). The high recitation rate for these two authors thus is in line with previous studies that reported higher recitation rates for prose (e.g. [1, 3, 4]).

In addition, our results revealed notable diachronic shifts in recitation patterns, as both recitation rate and pitch range varied depending on the decade of performance. Specifically, recitations in the 1980s were slower compared to earlier and later performances. From our results on articulation rate (time of articulation without pauses) and number of pauses (both not reported here), we know that both a more slowly articulation and the use of more pauses contribute to the lower recitation rate in that time. Additionally, we found that later recitations exhibited a wider pitch range, suggesting a more varied performance, while earlier recitations were faster and seemed to be more monotonous (with a smaller line-wise pitch range). However, there was a scarcity of data for certain decades compared to others, such as the 1980s, which only had recordings from three speakers. Nevertheless, an examination of Westphal's recordings, spanning from the 1960s to the 1990s, corroborated these diachronic trends, showing slower and more varied recitations in his later performances. To further explore this shift in recitation style, it will be necessary to enrich the corpus with more speakers and recitations across the 1970s, 1980s and the decades that were not covered in the current analyses due to scarcity.

Our preliminary results show how data-driven approaches reveal dependencies between different aspects of recitations (metadata, formal aspects, prosodic features) and can contextualize findings on individual speakers and poems. Additionally, the corpus analyses demonstrate that distant reading approaches can lead to interesting results regarding characteristics of particular authors (e.g. Heine) and speakers (e.g. Westphal) or recitation patterns related to a time period. For example, a particular way to read an author may be only detected using a lot of data. These findings are a promising starting point for close reading analyses (also in other disciplines like literary studies). For further studies, it seems fruitful to take more formal characteristics into account, such as, e.g. the absence or presence of rhyme and its influence on prosodic features.

We plan to use the insights from our corpus analyses to refine our speech synthesis models for poetry [18]. to incorporate specific styles, speakers, authors, or epochs (allowing us to synthesize a recitation to sound, for example, like a Goethe poem read by a female speaker in the 1950s). These models will allow us to investigate the aesthetic effects and functions of the investigated features in controlled perception studies.

#### 5. Acknowledgements

This research is supported by funding from the German Ministry for Education and Research (BMBF) for the »textklang« project.

## 6. References

- [1] P. Byers, “A formula for poetic intonation,” *Poetic*, vol. 8, pp. 367–380, 1979.
- [2] T. Barney, “Readers as text processors and performers: A new formula for poetic intonation,” *Discourse Processes*, vol. 28, no. 2, pp. 155–167, 1999.
- [3] S. Blohm, S. Versace, S. Methner, V. Wagner, M. Schlesewsky, and W. Menninghaus, “Reading poetry and prose: Eye movements and acoustic evidence,” *Discourse Processes*, vol. 59, no. 3, pp. 159–183, 2022.
- [4] J. Bröggelwirth, *Ein rhythmisch-prosodisches Modell lyrischen Sprechstils. Dissertation*, Universitäts- und Landesbibliothek Bonn, 2007.
- [5] P. A. Barbosa, “The dance of pauses in poetry declamation,” *Languages*, vol. 8, no. 1, 2023. [Online]. Available: <https://www.mdpi.com/2226-471X/8/1/76>
- [6] P. Wagner and S. Betz, “Effects of meter, genre and experience on pausing, lengthening and prosodic phrasing in german poetry reading,” in *Proceedings of Interspeech 2023*, Dublin, Ireland, 2023.
- [7] K. Schweitzer, K. Eckart, M. Gärtner, A. Falenska, A. Riester, I. Rösiger, A. Schweitzer, S. Stehwiem, and J. Kuhn, “German radio interviews: The GRAIN release of the SFB732 silver standard collection,” in *Proceedings of the 11th edition of the Language Resources and Evaluation Conference (LREC 2018)*, Miyazaki, Japan, 2018. [Online]. Available: <https://aclanthology.org/L18-1457>
- [8] N. Schauffler, T. Bernhart, A. Blessing, G. Eschenbach, M. Gärtner, K. Jung, A. Kinder, J. Koch, S. Richter, G. Viehhauser, T. Vu, L. Wesemann, and J. Kuhn, “>textklang<—towards a multi-modal exploration platform for German poetry,” in *Proceedings of the 13th edition of the Language Resources and Evaluation Conference (LREC 2022)*, Marseille, France, 2022, pp. 5345–5355. [Online]. Available: <https://aclanthology.org/2022.lrec-1.572>
- [9] S. Rapp, “Automatic phonemic transcription and linguistic annotation from known text with Hidden Markov models – An aligner for German,” in *Proc. of ELSNET Goes East and IMACS Workshop “Integration of Language and Speech in Academia and Industry” (Russia)*, 1995.
- [10] A. W. Black, “The Festival speech synthesis system,” <https://www.cstr.ed.ac.uk/projects/festival/>, November 1997.
- [11] Festival, “Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart. IMS German Festival home page,” Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, 2010. [Online]. Available: [www.ims.uni-stuttgart.de/phonetik/synthesis](http://www.ims.uni-stuttgart.de/phonetik/synthesis)
- [12] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2022. [Online]. Available: <https://www.R-project.org/>
- [13] D. Bates, M. Mächler, B. Bolker, and S. Walker, “Fitting linear mixed-effects models using lme4,” *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.
- [14] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, “lmerTest Package: Tests in Linear Mixed Effects Models,” *Journal of Statistical Software*, vol. 82, no. 13, p. 1–26, 2017. [Online]. Available: <https://www.jstatsoft.org/index.php/jss/article/view/v082i13>
- [15] R. V. Lenth, *emmeans: Estimated Marginal Means, aka Least-Squares Means*, 2022, R package version 1.8.3. [Online]. Available: <https://CRAN.R-project.org/package=emmeans>
- [16] J. A. Long, *jtools: Analysis and Presentation of Social Scientific Data*, 2022, R package version 2.2.0. [Online]. Available: <https://cran.r-project.org/package=jtools>
- [17] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. [Online]. Available: <https://ggplot2.tidyverse.org>
- [18] J. Koch, F. Lux, N. Schauffler, T. Bernhart, F. Dieterle, J. Kuhn, S. Richter, G. Viehhauser, and T. Vu, “PoeticTTS – controllable poetry reading for literary studies,” in *Proceedings of Interspeech 2022*, Incheon, Korea, 2022. [Online]. Available: <https://arxiv.org/abs/2207.05549>