Prosodic realisation of enjambment in recitations of German poetry

Nadja Schauffler\textsuperscript{1}, Fabian Schubo\textsuperscript{2}, Toni Bernhart\textsuperscript{3}, Gunilla Eschenbach\textsuperscript{4}, Julia Koch\textsuperscript{1}, Sandra Richter\textsuperscript{4}, Gabriel Viehhauser\textsuperscript{3}, Thang Vu\textsuperscript{1}, Lorenz Wesemann\textsuperscript{4}, Jonas Kuhn\textsuperscript{1}

\textsuperscript{1}Institute for Natural Language Processing, University of Stuttgart, Germany
\textsuperscript{2}Institute of English Linguistics, University of Stuttgart, Germany
\textsuperscript{3}Institute for Literary Studies/Digital Humanities, University of Stuttgart, Germany
\textsuperscript{4}German Literature Archive (DLA), Marbach, Germany

nadja.schauffler@ims.uni-stuttgart.de

Abstract

A salient feature of poetry is the organisation in lines and stanzas. Versification thus serves as a structural (and aesthetic) layer that either conforms to syntactic units or breaks them up. If line breaks disrupt syntactic units, we speak of enjambment - the line suggests a pause while the syntactic information continues. Since prosodic boundaries typically reflect syntactic boundaries, the question is which information is marked - the line or the syntactic unit. In a preliminary study with one professional speaker of German, we investigate how line breaks are prosodically realised in recitations of twenty poems by Friedrich Hölderlin. We compare cases of enjambment to line breaks without enjambment and look at lengthening of the line-final segment, frequency and duration of pauses and F\textsubscript{o} reset across the line break, which are typical cues used for the prosodic marking of phrase boundaries in speech. We found that while pauses and F\textsubscript{o} reset are tuned down in cases of enjambment, the line break is still prosodically marked by means of final lengthening. This preliminary result supports the idea that a speaker can convey both the syntactic continuity as well as the versification of the poem by strengthening different cues.

Index Terms: prosodic phrasing, poetry, enjambment, German

1. Introduction

This paper presents a baseline study from the »textklang« project aimed at exploring the interplay of the textual and the sound dimension of poetry – specifically German lyric poetry from the Romantic era. Over the course of the project, we will develop a methodological toolbox for examining an archive of recordings of poetry recitations and vocal performances of songs setting Romantic poetry to music. By combining prosodic analysis with text-analytical tools from corpus linguistics and computational literary studies, scholars are put in a position to formulate and test hypotheses regarding systematic patterns in the relationship between prosodic and textual form and aspects of interpretation. At a later stage, perception experiments employing controlled (re-)synthesized variants of the original recordings will add a further empirical dimension.

The present study plays an important role as a starting point for further method development. The prosodic characteristics of some particular recitation of a poem depend on a wealth of factors, ranging from individual choices by the performer, via styles established in some school of recitation to formal characteristics of the poem itself in its context of interpretation. The phenomenon of enjambment with its clear-cut anchoring in linguistic form on the one hand and poetic form on the other lends itself as a basis for calibrating some core dimensions of study and subsequently adding further factors. The study is based on a homogeneous, but also very characteristic dataset of recitations of twenty Friedrich Hölderlin poems by a single speaker, Hanns Zischler, who is a professional actor.

1.1. Enjambment in poetry

An enjambment occurs if the line end disrupts a syntactic unit so that versification suggests a prosodic boundary where we would not expect one from a syntactic perspective.

There are various ways to deal with this conflict as a reader or speaker - either conforming to the boundary as suggested by the syntax and reading over the line end, or realising a boundary at the line end as suggested by the poems’ written presentation thus disrupting the syntactic coherence.\textsuperscript{[1, 2]} suggest that there is a third possibility for a speaker, a so-called ‘rhythmical performance’. In such a recitation the speaker is said to prosodically convey versification on the one hand and the syntactic structure on the other hand by using both cues for continuation and cues for discontinuation simultaneously.

Speakers may, of course, be very different in their choice of performance, depending for example on the cultural background or their professional speaker education: Stefan George’s followers, for example, emphasized enjambment according to George’s own reading practice\textsuperscript{[3, pp. 704–705]; and [4]}, for example, found that especially poets from the former GDR emphasised the enjambment. Other factors which may affect the realisation of enjambment lie within the poem itself, most prominently the degree of disruption - a softer break between clauses, a stronger break within a phrase, or even a break within a word\textsuperscript{[5, pp. 25–34]}.

Poems differ in the extent to which they incorporate enjambment. Norbert von Hellingrath distinguishes between two poetic principles: the austere harmony\textsuperscript{(harte Fügung)} with a high frequency of hard enjambments (these poems typically do not rhyme) and the smooth harmony\textsuperscript{(glatte Fügung)} in which the line break coincides with a syntactic unit and mirrors natural prosodic breaks\textsuperscript{[8, p.1]}. According to Dionysius of Halicarnassus to whom Hellingrath refers, the austere harmony is disrupt and unpredictable, the smooth harmony is fluent and predictable\textsuperscript{[8, chapt. 21–23]}. A third style is the intermediate style which Dionysius praises as the best\textsuperscript{[8, chapt. 24]}. Consequently, Dionysius favors a word order where poetry benefits from the freedom of prose. He recommends amongst others enjambment as a means to heighten the beauty of a poem (ibid. chapt. 26).

\textsuperscript{1}It has been furthermore distinguished between prospective enjambment, where the enjambment is apparent at the end of the line, and retrospective enjambments which are only identifiable after the line break in the new line\textsuperscript{[6] see also [7]}. 
In the current pilot study, we investigate realisations by one professional speaker and his recitations of poems by one German author as a starting point to a larger study on different speaker performances on the one hand and their effects in perception on the other hand. We wanted to know how this speaker performs cases of enjambment in these poems, more specifically whether he uses prosodic boundary cues to prosodically mark the line break.

1.2. Prosodic boundaries

Speakers use various phonetic correlates in the production of prosodic boundaries, including changes in segment duration, the insertion of pauses, and \( F_0 \)-register effects.

As for segment duration, it has been found across languages that the segments are produced with longer duration when they occur immediately preceding a boundary (e.g., [9, 10] for English, [9, 10, 11] for German). This lengthening effect is strongest on the elements of the phrase-final rhyme, but can also apply to earlier material (e.g. [9, 11, 12] for German). The amount of lengthening is correlated with boundary strength, so that relatively stronger boundaries involve a relatively longer duration of the final segments (e.g., [13]).

Another durational correlate of prosodic boundaries is the insertion of a silent pause. This is an optional boundary marker that occurs only at relatively strong boundaries (e.g., [14]). The duration of a pause is considered to reflect the relative boundary strength, so that longer pauses are associated with relatively stronger boundaries (e.g., [15]).

With respect to \( F_0 \)-register effects, it has been found that \( F_0 \) undergoes lowering throughout a prosodic phrase or utterance (e.g., [16]). This lowering effect is interrupted after a stronger prosodic boundary by resetting the upper limit of the \( F_0 \) register to a higher level (e.g., [17, 18]).

For German, it has been observed that stronger prosodic boundaries are more likely to occur with bundles of boundary markers [19]. Furthermore, it has been shown that relative boundary strength reflects syntactic bracketing in such a way that stronger syntactic boundaries are indicated by the implementation of stronger prosodic boundary cues (e.g., [20, 21, 22]).

In the current study, we investigated the presence of these cues at lines with enjambment compared to lines without enjambment. We hypothesised that line ends with enjambment are realised with stronger phrasing cues than syllables within the line but with weaker cues than line ends without enjambment. Specifically, we expected to find “line-final lengthening” as a cue for discontinuation to be used at lines with enjambment, while we expect pauses and \( F_0 \) reset not to be used to the same extent by the speaker in lines with enjambment compared to syntactically complete lines without enjambment, since these cues are typically found at stronger boundaries.

Frequency and length of pauses have been also investigated by [4] as a feature to identify and distinguish emphasised and unemphasised enjambment when the pause after the line end was shorter than pauses between words within the line, they classified it as unemphasised enjambment. By including further phrasing cues we may detect the prosodic marking of enjambment lines on a more subtle scale.

2. The data

The data analysed is a subset of a corpus comprising recitations of poems from the Romantic era which is currently being created. The subset comprises 20 poems by the German poet Friedrich Hölderlin which were recited by the actor Hanns Zischler for the sake of an exhibition.2 The recitations of these poems make up 43.65 minutes of speech. The poems vary in length from 9 lines (the shortest poem) to 160 lines (the longest poem) and 568 lines altogether (not counting the reproductions of the poems’ titles or of the author name).

3. Data processing

The textual data was automatically tokenized with the TreeTagger [24] and end of lines and end of stanzas were annotated.

The acoustic data was force-aligned for phone, syllable and word boundaries [25]. For each syllable in the data, PaIntE parameters were calculated [26, 27, 28]. PaIntE stands for “Parametrized Intonation Events” and presents a way to model the shape of the \( F_0 \) contour in the vicinity of intonation events. The model approximates a peak in the smoothed \( F_0 \) contour by means of six free parameters. These parameters depict phonetic cues on and around the accented syllable (e.g. peak height, peak alignment, amplitude of the rise). We used the parameter that corresponds to the height of the peak in our analysis (parameter \( d \)). Segment duration, and frequency and duration of pauses were extracted by means of the Festival [29] version of the University of Stuttgart [30] - a synthesis system for German.

Intonation events, in terms of GroB(S) labels [31] for pitch accents and boundary tones, were automatically annotated as described in [12]. In this procedure, intonation events are predicted by means of Random Forest classifiers on the bases of PaIntE parameters, normalised phone duration, lexical stress, syllable position, following silence, part of speech and punctuation.

4. Analyses

In order to investigate the realisation of enjambment by our speaker, we compared the poems’ lines with enjambment to the poems’ lines without enjambment. In lines with enjambment, the line break disrupts a syntactic unit. Since a punctuation mark signals either the completion of a sentence or a clause, we defined enjambment as the absence of a punctuation mark ( , : ?) at the end of the line. This corresponds to 202 lines out of the 568, i.e. 36% of all lines are with enjambment. See below for an example for a realisation of a stanza with three lines with enjambment and the last line without enjambment from Hölderlin’s poem Heidelberg.3

Quellen hattest du ihm, hattest dem Flüchtigen
Kühle Schatten geschenkt, und die Gestade sahn
All ihm nach, und es bette
Aus den Wellen ihr lieblich Bild.

4.1. Duration of the last segment

Since final lengthening is a prominent cue for the prosodic marking of phrase boundaries, we first investigated whether the

---

2Before 1800, Hölderlin prefers the smooth enjambment, in the poems after 1800, hard enjambment becomes the rule, including the disruption within a word (morphological enjambment) [5, p. 31]. Enjambments even interconnect stanzas and thus make the dissociation of syntactical and metric units complete [23, p. 262].

3Springs you had him, had the fleeting // given cool shadows, and the shores looked // all after him, and it quivered // out of the waves her lovely image. (own translation)
end of the line provoked a lengthening of the final segment. We calculated z-scores for the duration of the last segment of all syllables in the data by using phoneme-specific means and standard deviations in order to normalise for phone-inherent duration differences. Four segments had to be removed from the statistical analysis since their duration could not be determined.

4.1.1. Statistical analysis and result
Statistical analyses for this and the following features were performed in R 3.5.0 [32], using the function lmer from package lme4 [33]. We performed a linear mixed effects analysis with the normalised duration as dependent variable. As fixed factors we included an interaction between punctuation and position in line (with no punctuation and line end depicting the case of enjambment). In order to control for lengthening effects caused by lexical stress and/or pitch accenting, we also added presence of a pitch accent on the syllable and whether the syllable is lexically stressed as fixed factors. As random factors, we included intercepts for word and for poem. All factors were tested for their significance by means of likelihood ratio tests by comparing the model including the effect in question to the model without it.

Figure 1 gives an overview of the z-scored duration of a syllable’s last segment by its position in the line and by punctuation. Looking at last segments within the line (Intercept: $\beta=-0.27$, SE=0.05, $t=-5.20$), a punctuation mark significantly increases the duration to a z-score of 0.18 ($\beta=0.45$, SE=0.04, $t=12.78$). The effect of a punctuation mark is even stronger at the end of the line for which the model predicts a z-score of 0.56 ($\beta=0.46$, SE=0.09, $t=-5.26$). The line break without punctuation, which is the case of enjambment, increases the duration of the last segment to a similar extend to a z-score of 0.57 (compared to segments within the line without punctuation) ($\beta=0.87$, SE=0.07, $t=12.97$). We relevelled punctuation and position-in-line to confirm that the difference between line ends with punctuation mark and line breaks without punctuation mark is not significant ($\beta=0.01$, SE=0.08, $t=0.12$).

### 4.2. Frequency and duration of pauses
The next cue we investigated was whether the speaker realised a pause after the line break. Pauses are typically used to mark stronger phrase boundaries (see above), where the pause is expected to be longer the stronger the boundary.

4.2.1. Statistical analysis and result
We found that out of the 202 lines with enjambment, the speaker realised a silence after the line in 80 cases i.e. in 40% of the time. While of the 366 lines without enjambment, 326 are realised with a pause, i.e. in 89% of the time.

In order to investigate whether the presence of enjambment affects the duration of a possible pause, we fitted a linear mixed model with pause duration as dependent variable. As fixed factor we included punctuation, i.e. whether the line ends with an enjambment (no.punctuation) or with a punctuation mark (punctuation). As random factor we included an intercept for poem. A model comparison between the model with the factor punctuation to the model without punctuation revealed that this factor significantly improves the model ($\chi^2(1)=49.66$, $p<.001$). Figure 2 shows the effect as predicted by the model. When the line ends with a punctuation mark, the pause is significantly longer than after lines with enjambment ($\beta=0.48$, SE=0.06, $t=7.38$).

4.3. $F_0$ reset
The third parameter we analysed was whether there was a considerable $F_0$reset in the new line as would be expected after a more prominent phrase boundary. Therefore, we used the PaInTE parameter $d$, which describes the height of the peak in a three-syllable window. We extracted this parameter on the last syllable in the line that has been automatically annotated with a pitch accent and subtracted the PaInTE parameter $d$ of the first syllable carrying an automatically annotated pitch accent in the next line. A negative result points toward a $F_0$ reset which we expect for larger phrase breaks, a positive number suggests $F_0$ declination as expected within a phrase. In 140 cases an accent could not be automatically detected in either the current line or in the next line so that these cases were discarded for this analysis. This affected 66 lines with enjambment and 74 lines with

![Figure 1: The duration of a syllable’s last segment as z-score by the presence of a punctuation mark and by the position in the line. The case of ‘no punctuation’ at ‘line.end’ (the third box) depicts the case of enjambment.](image1)

![Figure 2: The duration of the silent pause after the line in lines with no punctuation (= enjambment, left box) compared to lines with punctuation mark (right box).](image2)
punctuation mark.

4.3.1. Statistical analysis and result

We fitted a linear mixed model with the difference between the last peak in the line and the first peak in the next line in Hertz as dependent variable. As fixed factor we included punctuation, i.e. whether or not the line ended with a punctuation mark (no punctuation = enjambment). As random factors, we included intercepts for the types of predicted pitch accents as a sequenz (PA-line1.PAline2), and for poem.

The factor punctuation significantly improved the model fit as established by means of a likelihood ration test ($\chi^2(1)=50.25, p<.001$). See Figure 3 for an overview of the difference between the peaks across the line break. Negative difference suggests a F$_0$ reset, while positive differences entail a down-drift across the line break. At the presence of a punctuation mark (right box), the difference between the peak is predominantly negative ($\beta=-28.14, SE=3.52, t=-7.99$), suggesting a F$_0$ reset.

Figure 3: The difference in Hz between the height of the peak (PaIntE parameter d) on the last predicted accent in the line and the height of the peak on the first predicted accent in the next line by the presence of a punctuation mark. The case with no punctuation park (left panel) depicts enjambment.

5. Discussion and conclusion

We investigated whether we find typical boundary cues at line breaks with enjambment in twenty German poems spoken by one professional speaker. We hypothesised that we will find lower-threshold cues (final lengthening) but not cues typically associated with stronger boundaries, such as silent pauses and F$_0$ reset.

We found that the final segment in a line indeed is reliably lengthened irrespective of whether the line disrupts syntactic units as in the case of enjambment, or whether it completes a syntactic unit. We can say that the speaker prosodically marked the formal presentation of the single lines by means of line-final lengthening. The other two cues which we investigated, namely frequency and duration of pauses and F$_0$ reset was not used in the same extend as for complete phrases at line breaks: The speaker uses considerably fewer silent pauses after lines with enjambment and if he realises a pause, it is significantly shorter than in the case of lines without enjambment. Similarly, we found significantly more often F$_0$ reset in lines without enjambment, a cue expected at stronger prosodic boundaries. Lines with enjambment on the other hand were more often realised with F$_0$ downdrift across lines, a phenomenon we usually find within prosodic phrases.

Our results support the assumption that speakers can use cues for discontinuation (final lengthening) and cues for continuation (no pause or short pause, and no F$_0$ reset) to simultaneously convey the formal end of the line and the completeness of syntactic structure. Additionally, our results show that enjambments without a silent pause after the line (unemphasised enjambment) in [4]) may still be prosodically emphasised by means of final lengthening.

Our results also support findings that phrasing cues are used in bundles depending on the strength of the boundary [19]. This suggests that there may be a difference in the prosodic marking of lines between different kind of enjambments with ‘soft’ enjambments provoking stronger phrasing cues than ‘hard’ enjambments where the continuation constraint may be especially strong.

6. Outlook

As we laid out in the introduction, the aim of the present study was to establish a baseline for further work in the »textklang« project, both conceptually and in terms of method and tool development. Next steps in our prosodic research will attend to a more fine-grained analysis of different kinds of enjambment. At the same time, we will constantly extend the corpus to include poems from a range of poets from the Romantic era as well as different speakers performing the recitation, which will allow us to compare different speakers’ strategies. The German Literature Archive Marbach has a rich collection of recordings, reaching back to the 1920s, including a broad variety of speakers.

Building on prior methodological work supporting corpus analysis with a combination of text-oriented and speech-oriented models [34], the »textklang« project is developing a multi-modal corpus exploration platform. The scope of analysis departs from linguistic aspects of prosody (as exemplified by the baseline study reported in this paper), but combines this with theories and insights from literary studies, e.g. about the status that the Romantic era attributed to sound. The combination of models and work practices from distinct disciplines studying aspects of language, speech and texts in a given context of reception is an extension of an approach developed over several years in interdisciplinary work at the Center for Reflected Text Analytics CRETA [35, 36, 37]. The sound level of literary texts has not been addressed so far. The »textklang« project will not only support analytical work on existing corpora and collections of recordings, but also hypothesis testing via controlled perception experiments: the exploration framework will allow experimenters to synthesize variants of recitations of poems by systematically manipulating prosodic parameters by ways of testing some underlying hypothesis, which can be informed by considerations from linguistic theory as much as literary studies.

7. Acknowledgements

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