THE EFFECT OF PROSE AND POETRY TEXT LAYOUT ON THE ACOUSTICS OF READING ALOUD

Filip Nenadić[1], Benjamin V. Tucker[2,3]

[1]Singidunum University; [2]Northern Arizona University; [3]University of Alberta
filip.nenadic@fmk.edu.rs, benjamin.tucker@nau.edu

ABSTRACT

Previous studies show that the same text is cognitively processed and interpreted differently based on the layout it is presented in – differentiating, for example, prose from poetry. Differences in how these two distinct literary genres are approached could influence the acoustic properties of their oral reading. A recent study by Blohm et al. [1] showed that the same artificially constructed short text presented in a poetic layout (with verses) was read slower and with more frequent, yet shorter pauses. In the present study, we investigate the effect of prose versus poetry text layout on oral reading of ecologically valid short texts written and formatted by published poets. We attempt to replicate the speech production findings from Blohm et al., yet find partly different results. We discuss how these disparate outcomes in oral production of prose and poetry may depend on text characteristics (rhythm and meter, difficulty) and reader characteristics.

Keywords: poetry, genre, oral reading, speech acoustics, prosody

1. INTRODUCTION

Characteristics of speech production vary depending on genre and register. Recent studies of oral reading have examined differences in the way in which poetry and prose is read aloud. The genre-specific hypothesis of reading claims that the first step in any reading is determining the text's genre and adjusting cognitive processing depending on the genre of the text [2, 3, 4]. Studies show that changing the text layout from prose to verses changes readers' genre decision, processing strategies, interpretation of a given text, speech acoustics in oral production [1, 5, 6, 7].

Where speech acoustics of oral reading is concerned, in a case study in Swedish, Fant [8] found that two poems were read more slowly and with a more stable tempo, with higher intensity, and with a higher and less varying F0 when compared to their reformatted prose forms and a separate novel excerpt. Similarly, Tsur [9] focused on case-studies of expert recitations of Shakespeare and Milton, stressing the importance of line breaks and meter on performance.

More recently, Blohm et al. [1] conducted an experimental study in which they constructed short texts in German. The texts consisted of two sentences and the authors altered their layout from prose to poetry. Their participants were instructed to read the texts aloud and their eye-movements were tracked. The participants read texts presented in a poetic layout more slowly as both their articulation rate was slower and the total time spent in silent pauses was longer. The pauses were more frequent and on average shorter when reading a text presented in verse. These differences were also reflected in the recorded eye-movements.

Furthermore, the constructed texts featured a rhythmic pattern in which strong and weak syllables alternated. The authors focused on critical regions consisting of two function words at the beginning of a sentence (which could either be the first or the second sentence in the short text), followed by a post-critical region content word. They observed that the duration ratio between the strong and the weak syllables in the critical region was larger when a text was read in its poetry versus prose format. Such local differences were not noted with regard to ratios of strong/weak syllable intensity or pitch. In the post-critical region, both strong and weak syllables were longer in the poetic layout.

In the present paper, we expand on this line of research by investigating differences in oral readings of prose and poetry in English. Our participants have a similar background to those tested by Blohm et al. [1] and we use a number of the same acoustic measures to investigate their productions of prose and poetry stimuli. However, the key difference is that in our study we did not construct texts to have a particular rhythmic pattern. Instead, we used original short texts written and formatted by young published authors for the purposes of this study. Therefore, the central research question of the present paper is whether text layout (prose versus

poetry) has an impact on oral reading of literary texts even when there is no meter or rhythmic pattern, that is, when free verse is used.

2. METHODS

2.1. Participants

Thirty-six native speakers of English participated in the study (11 male; ages 18-32). All of the participants were recruited from the University of Alberta community. The majority were undergraduate students enrolled in an introductory linguistics course who received course credit for their participation. Participants were recruited as non-experts in literary reading.

2.2. Stimuli

The stimuli were 20 original short texts (word count M = 59.4, SD = 9.13) written by seven young published authors for the purposes of the present study. The authors were instructed to write texts that they deemed acceptable as both short prose and poetry. These texts were further formatted in four layouts by their authors: double-aligned (justified) and left-aligned prose, verses in a single stanza, and verses in two or more stanzas, for a total of 80 different stimuli (see Figure 1 as an example). Three additional texts written by some of the same authors were used as practice stimuli.

Since no other restrictions besides adequacy in both the prose and poetic format were imposed in text creation, the stimuli did not rhyme or have a strict metric or rhythmic structure. Effectively, the poetic layouts can therefore be considered free verse [10, 11].

2.3. Procedure

The participants were given the task to read aloud the texts appearing on a computer screen. Their instruction was to read the texts the way they would read them aloud in company of others or if, for example, a teacher were to ask them to read aloud for the rest of the class. Participants' voice and eye-movements were recorded during reading (although we do not analyze the eye-movement data in the present report).

The participants first read three practice texts. They proceeded with reading the 20 texts, five in each of the four text layouts. The layout in which a certain text would appear was counterbalanced across participants using a Latin square design. The text presentation order was randomized for each participant; different text layouts were not placed in separate blocks.

A short break was offered after 10 texts were read. Upon reading all 20 texts and another break, the participants read the same texts in the same layouts again, but in a new randomized order (although we only analyze the first pass in the present report).

2.4. Data Analysis

One reading of a certain text by one participant constituted a single recording. These recordings were forced aligned using the Penn Forced Aligner [12]. All of the aligned recordings were manually inspected and all recordings that contained an error in reading were excluded.

We could not replicate the analysis Blohm et al. [1] on the local level because our stimuli do not feature the same strong/weak syllable pattern. Instead, we focused on the acoustic measures calculated on the global level. We considered seven global measures of oral reading and calculated them for each of the retained recordings. Two of these seven measures were standard deviation of intensity and fundamental frequency. Both the intensity and the fundamental frequency values were automatically extracted for vowels using a Praat script [13]. The other five global measures were identical to those considered by Blohm et al. [1]:

- **speech rate** – number of syllables divided by the total reading time (pauses included)
- **articulation rate** – number of syllables divided by the total time spent articulating (i.e., with silent pauses excluded)
- **pause rate** – number of syllables divided by number of pauses
- **pause duration** – summed duration of all silent pauses
- **pause proportion** – pause duration divided by the total reading time

Manual inspection of the recordings showed that 39% had at least one reading error. These recordings were excluded for two reasons. First, automatic alignment cannot function properly if changes are introduced in comparison to the expected text. Second, we cannot know what impact a reading error may have on overall reading performance and, specifically, the acoustic measures we estimated. We further excluded four participants that had only six or fewer recordings without errors remaining. The final number of retained recordings was 419 and a chi-square test showed that all four text layouts were equally present among the retained recordings ($\chi^2 = 0.33$, $df = 3$, $p = .95$).

Each of the described measures of oral reading

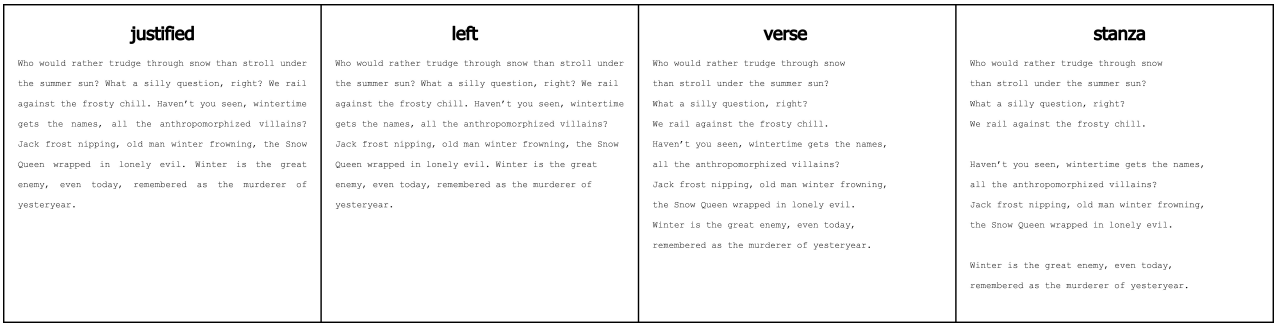

Figure 1: Sample text-stimulus in four layouts.

was used as a dependent variable in a separate linear mixed effects model. Model fitting was performed in R [14] using the lme4 package [15]. We created two sets of models. In one set, the independent variable was text layout with four levels (justified prose, left-aligned prose, verses, verses in stanzas). Post-hoc analyses in these models were performed using the multcomp package [16]. In the second set of models, we only considered the distinction between prose (conflating justified and left-aligned prose) and poetry (conflating verses in one stanza and verses in multiple stanzas). Random intercepts per subject and per text were included in all models. The independent variable (one of the four text layouts or the prose versus poetry distinction) was considered significant if it significantly improved model fit when it was added to a baseline model that included random intercepts only. Model criticism and trimming was performed for all models; model trimming never excluded more than 2.5% of data.

3. RESULTS

We found a significant contribution to model fit in four cases. In the case of speech rate, a significant contribution of the independent variable was registered both when all four layouts (justified prose, left-aligned prose, verses, and stanzas) were considered separately and when a simple prose versus poetry distinction was made. Since it is more informative, we will report on the model that distinguished between the four layouts. We found that texts formatted as left-aligned prose were produced with a higher speech rate than texts formatted as verses ($\beta = 0.35$, $SE = 0.12$, $z = 2.85$) and stanzas ($\beta = 0.40$, $SE = 0.13$, $t = 3.15$). Other pairwise comparisons were not significant (Figure 2).

We also recorded a significant effect of text layout on articulation rate and proportion of pauses, but these effects reached significant levels only when

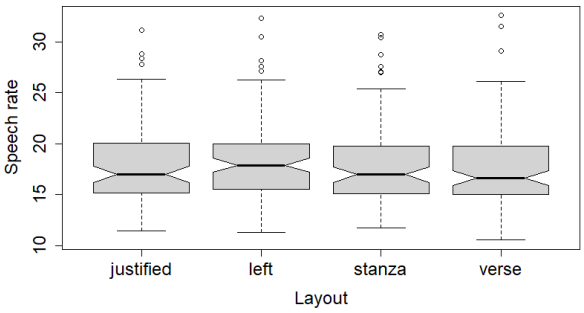

Figure 2

Figure 3: Boxplot of differences in speech rate between the four textual layouts.

we considered the two prose (justified and left-aligned) and the two poetry layouts (verses and stanzas) jointly (Figure 4). Lower articulation rate was registered for poetry ($\beta = -0.20$, $SE = 0.08$, $t = -2.41$). In other words, when they articulated, participants articulated fewer syllables per unit of time. Additionally, participants had a higher proportion of pauses when they read poetry, meaning that they spent more of the total speaking time being quiet ($\beta = 0.005$, $SE = 0.002$, $t = 2.26$).

4. DISCUSSION

The goal of the present paper was to test whether there are differences in acoustic measures of oral productions of the same short literary texts when formatted as prose versus poetry. Studies such as this one serve to inform us about (1) the reading strategies, cognitive processing, and interpretation of different (literary) genres and (2) variability in speech production dependent on the genre/register and purpose of speaking.

The number of potential indicators of performance that can be extracted from a recording of oral reading

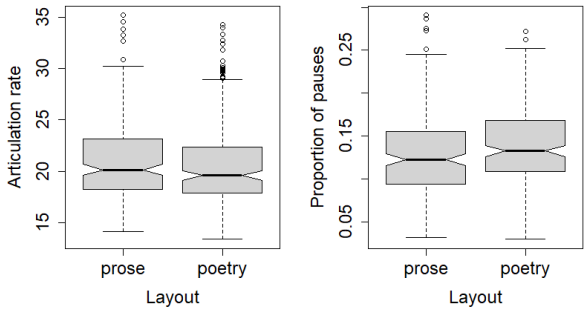

Figure 4

Figure 5: Boxplot of differences in articulation rate and proportion of pauses between prose and poetry.

of a brief text is very large. Different researchers may opt for different operationalizations and these researcher degrees of freedom may have an impact on the results of the study [17]. To avoid this possibility, we utilized the same global measures of oral reading as Blohm et al. [1]. Additionally, we included estimates of standard deviation of fundamental frequency and intensity on the global level, which Blohm et al. [1] analyzed in their local, critical regions.

We find differences in the way the same texts are read when they are presented in the prose versus poetry layout. Our participants generally spent more time reading poetry, as their speech rate was slower. Both lower articulation rate (fewer syllables produced per unit of time when articulating) and higher proportion of pauses (more total reading time spent silent) contributed to this difference in speed. However, we do not find evidence of more frequent pauses, nor differences in pause duration.

Since the participants tested in Blohm et al. [1] and our studies had similar backgrounds, another possible reason could be the differences in the way the stimuli in these two studies were created and, consequently, their characteristics. Our stimuli were written by young published authors for the purposes of the study with instruction to write texts that they believe could be fit as both short prose and a short poem, yielding texts that did not feature a consistent metric or rhythmic structure.

Although periodical repetition of rhythmic patterns is considered a frequent formal constraint of conventional regulated verse, it is not necessarily present in all poetry [10, 11]. Still, the removal of these characteristics may eliminate cues used for expressive poetry reading; free verse is often described as similar to conversational speech. Perhaps our readers found no place for an expressive pause or the need to make longer, more dramatic pauses. Previous studies show that reader appreciation and even ratings of the text as a poem ('poeticity' ratings) dwindle when meter and rhyme were deteriorated or removed from poems when they are expected to be found [18, 19, 20]. If this is true, then future studies should examine which acoustic aspects of expressive oral reading are changed depending on presence or absence of certain text characteristics or poetic devices.

Another possibility is that the stimuli we employed were too long and/or too difficult for our participants. We registered reading errors in almost 40% of recordings, much more than the 11% recorded by Blohm et al. [1]. Studies that tested children's oral reading show that decoding difficulties impact reading prosody by increasing the number and duration of (unexpected) pauses and making the F0 prosodic profile less similar to those registered in proficient, adult speakers [21, 22]. Difficult texts seem to make it more difficult for readers to maintain focus on comprehension, increasing mind wandering [23]; if a reader is allocating cognitive control and resources to decoding, it may be difficult to simultaneously provide an expressive oral reading performance.

This interpretation should be considered with caution, however. It could be that the higher number of errors we record is simply a consequence of a longer text (and therefore more opportunity to make a mistake), not a consequence of text difficulty. The Flesh reading ease score of our texts ranged between 42.4 and 98.2 ($M = 76.19$, $SD = 14.27$), with lower values indicating that a text is more difficult to read. In comparison, text difficulty in the study conducted by Blohm et al. [1] ranged between 40 and 87 ($M = 65.33$, $SD = 8.87$).

Despite these stimuli differences, we cannot be sure that they are the (sole) cause behind partly disparate results. The two studies tested speakers of different languages which may be behind the somewhat disparate results. Cross-linguistic investigation is necessary to expand our understanding of topic. Perhaps the necessary exclusion of recordings reduced our statistical power to capture all the fine differences in relevant acoustic measures, as these effects are often subtle. In the present report, we only considered the first reading of the texts from each participant. A pending analysis of their second pass should lead to an increase in the statistical power, but also potentially yield fewer errors and more expressive reading after the practice participants had with the stimuli.

5. REFERENCES

[1] 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.

[2] D. Hanauer, "The genre-specific hypothesis of reading: Reading poetry and encyclopedic items," *Poetics*, vol. 26, no. 2, pp. 63–80, 1998.

[3] R. A. Zwaan, "Some parameters of literary and news comprehension: Effects of discourse-type perspective on reading rate and surface structure representation," *Poetics*, vol. 20, no. 2, pp. 139–156, 1991.

[4] ——, "Effect of genre expectations on text comprehension," *Learning, Memory, and Cognition*, vol. 20, no. 4, pp. 139–156, 1994.

[5] M. Fechino, A. M. Jacobs, and J. Lüdtke, "Following in jakobson and lévi-strauss' footsteps: A neurocognitive poetics investigation of eye movements during the reading of baudelaire's 'les chats'," *Journal of Eye Movement Research*, vol. 13, no. 3, 2020.

[6] D. Hanauer, "Integration of phonetic and graphic features in poetic text categorization judgments," vol. 23, no. 5, pp. 363–380, 1996.

[7] J. Peskin, "The genre of poetry: Secondary school students' conventional expectations and interpretative operations," *English in Education*, vol. 41, no. 3, pp. 20–36, 2007.

[8] G. Fant, *Speech acoustics and phonetics*. Dordrecht: Kluwer Academic, 2004.

[9] R. Tsur, "Poetry reading–rhythmical performance: Triple-encoding, and voice quality: Six case studies," *Thinking Verse*, vol. 2, pp. 88–111, 2012.

[10] N. Fabb, "There is no psychological limit on the duration of metrical lines in performance: Against turner and pöppel," *International Journal of Literary Linguistics*, vol. 2, no. 1, pp. 1–29, 2013.

[11] ——, *What is Poetry?: Language and Memory in the Poems of the World*. Cambridge University Press, 2015.

[12] J. Yuan and M. Liberman, "Speaker identification on the SCOTUS corpus," *Proceedings of Acoustics '08*, pp. 5687–5690, 2008.

[13] P. Boersma and D. Weenink, "Praat, a system for doing phonetics by computer," 2019. [Online]. Available: www.praat.org

[14] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2020. [Online]. Available: https://www.R-project.org/

[15] 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.

[16] T. Hothorn, F. Bretz, and P. Westfall, "Simultaneous inference in general parametric models," *Biometrical Journal*, vol. 50, no. 3, pp. 346–363, 2008.

[17] S. Coretta, J. V. Casillas, and T. B. Roettger, "Multidimensional signals and analytic flexibility: Estimating degrees of freedom in human speech analyses," Nov 2022. [Online]. Available: psyarxiv.com/q8t2k

[18] D. Hanauer, "Reading poetry: An empirical investigation of formalist, stylistic, and conventionalist claims," *Poetics Today*, vol. 19, no. 4, pp. 565–580, 1998.

[19] C. Obermeier, W. Menninghaus, M. Von Koppenfels, T. Raettig, M. Schmidt-Kassow, S. Otterbein, and S. A. Kotz, "Aesthetic and emotional effects of meter and rhyme in poetry," *Frontiers in Psychology*, vol. 4, p. 10, 2013.

[20] W. van Peer, "The measurement of metre: Its cognitive and affective functions," *Poetics*, vol. 19, no. 3, pp. 259–275, 1990.

[21] R. G. Benjamin and P. J. Schwanenflugel, "Text complexity and oral reading prosody in young readers," *Reading Research Quarterly*, vol. 45, no. 4, pp. 388–404, 2010.

[22] P. J. Schwanenflugel, A. M. Hamilton, M. R. Kuhn, J. M. Wisenbaker, and S. A. Stahl, "Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers," *Journal of Educational Psychology*, vol. 96, no. 1, pp. 119–129, 2004.

[23] S. Feng, S. D'Mello, and A. C. Graesser, "Mind wandering while reading easy and difficult texts," *Psychonomic Bulletin & Review*, vol. 20, no. 3, pp. 586–592, 2013.