PROSE OR POETRY? A PERCEPTIVE PHONETIC STUDY

Valentina Colonna[1], Antonio Romano[2]

[1]University of Granada (Spain), Department of General Linguistics and Literary Theory
[2]University of Turin (Italy), Department of Foreign Languages, Literatures and Modern Cultures
valentinacolonna@ugr.es, antonio.romano@unito.it

ABSTRACT

In their spoken form, poetry and prose distinguish in many and acknowledged aspects, but are they discernible at a perceptual level by their prosody? This topic received very little attention. This paper aims to present the results of a perceptive Phonetics study based on the recognition, by a group of students, of prose and poetry among manipulated delexicalised stimuli. Taking into account a selection of variables, the perceptive approach, combined with statistical analysis, enabled us to evaluate also the influence of the textual construction, the speaker's vocal type, and the audio manipulation. Among the main outcomes, we observed poetry recognition in its "metrical" reading and through characterizing traits, such as longer pauses and a slower speech rate, connected to a specific use of melody. This task has proved to be a useful tool for deepening the acoustic dimension of the literary text, putting the focus on the listener's perception.

Keywords: Perceptive Phonetics; Experimental Task; Poetry; Prose; Voices of Italian Poets

1. INTRODUCTION

Poetry and Prose diverge in many aspects that have effects also on the prosodic dimension: first of all, silence and layout, length, density, rhythm, and melody. However, although the rhythmic issue of both has been addressed from a theoretical point of view, Linguistics works concerning the recognition of poetry and prose on a perceptual level are still little considered[1].

Among the most representative, we can mention Fónagy's [1] phonological approach, Bolinger's [2] studies, Bröggelwirth's [3] work in the German context, Wagner's [4] approach, which focused on metrical feet, and MacArthur and Miller's [5], focusing on variables in spontaneous speech and poetry readings.

While studies showed poetry and prose readings' distinctions in the prosodic production, on a perceptual level, in the German context [3] [6], the issue has not yet been addressed for Italian, of which this paper represents an initial proposal of study.

This work takes into consideration previous linguistic studies aimed at other forms of speech. The main reference works include 't Hart et al. [7] and Romano [8] [9]: in this study, a particular reference will be to the AMPER project methodology [10]. Further literature references were Moulines & Verhelst [11], Van Bezooijen & Gooskens [12], and Knoll *et al.* [13].

2. THE STUDY

This study has been conducted as part of the VIP-*Voices of Italian Poets* project [14] and consists of the analysis of a perceptive Phonetics task submitted to the attention of a group of native Italian students and composed of manipulated audio data, which we go to introduce.

2.1. Data

The acoustic data forming part of the experimental task consist of recordings by an actress-announcer (woman, an actor (man), and a poetess (woman)). The three speakers read two poems by two different authors of the Italian 20th century: one by Giorgio Caproni (*Alba*) and one by Giuseppe Ungaretti (*Sirene*).

The first composition is in fixed metrics (sonnet), but with a language and style closer to our time; the second poem is in free metrics (mainly with hendecasyllables and septenaries), with more conservative language and style. Although with overall short lengths, the poems are very different in form and content. The selection criteria also depend, besides the existence of the original recording by the poet (see Colonna [15]), on the hypothesis of stylistic reading changes depending on the theme-form combination (rhythmic-metric) (see Bologna [16]) and the poem's rhythmical, syntactical, and rhetorical issue.

The choice of different types of speakers was made considering the need to test different prosodic styles and approaches to the poetry reading, to assess their influence on perception.

The speakers were asked to read firstly "metrically" (see [15]), following the layout of the poem as closely as possible, respecting the verse as the unit of measurement, and, then, to read "syntactically", following the punctuation and syntaxis, the same text transposed into prose.

The recordings were made in the "Arturo Genre" Laboratory of Experimental Phonetics of the University of Turin. Each file has been sampled at 16 kHz, and agreed to let their voice be used for this study under the GDPR.

2.2. Methodology

Poetry and prose presented clear characteristic traits in all three speakers: lower Fluency and Speech Rate, internal melodic variety, and greater use of pauses, broader overall, in the first case; higher Fluency and Speech Rate, higher internal melodic variety and less use of pauses, shorter overall, in the second case.

Overall, we collected 6 recordings, including prose/poetry prosodic readings of every poem (2) among the 3 speakers. For each recording, a selected portion of a maximum of 20 s has been detected and annotated at a vowel level through PRAAT program. To compare the degree of perception between different types of alterations, we manipulated the data. More in detail, we used a script to extract some data (duration, f_0, and intensity), then manually corrected and manipulated in 4 ways: energy normalisation; flattening of f_0 average and energy normalisation; flattening of durations and energy normalisation; further flattening of durations (shorter) and energy normalisation. Later, these data were entered into MATLAB software which, through the AMPER routines [17], produced 4 synthetic audios per recording (delexicalised stimuli), as well as one synthetic audio without further manipulation.

Fig. 1 and Fig. 2 show two examples of manipulation (fixed intensity and fixed f_0) in announcer-actress' prose reading.

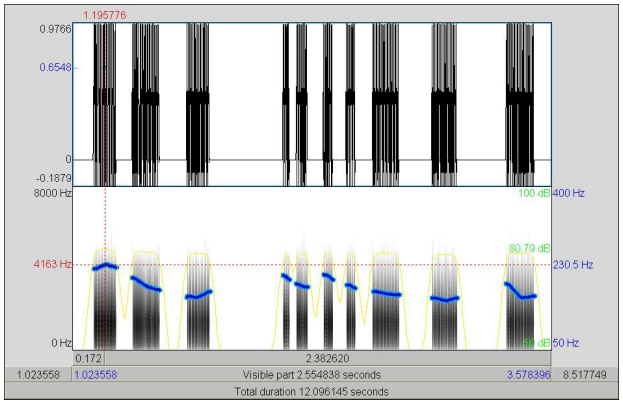

Figure 1: PRAAT window of synthetic voice stimulus (fixed intensity).

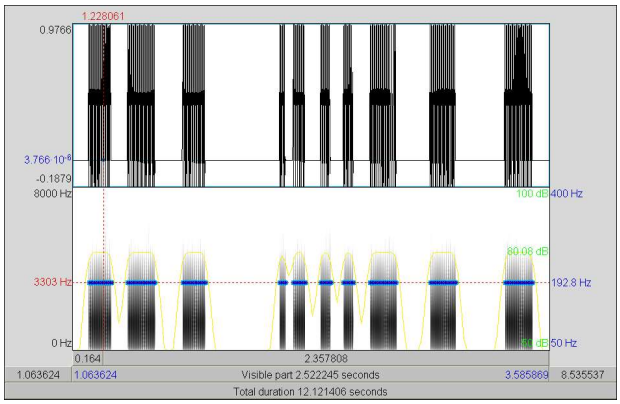

Figure 2: Praat window of synthetic voice stimulus (fixed f_0).

Before the creation and submission of the final task (by PRAAT), we create and test a training test (identification test) aimed at the recognition of prose or poetry (in random order). The task we will present in this paper is a discrimination test composed of 32 random input pairs of stimuli, each proposing one prose and one poetry, with equal handling characteristics and the same speaker. The listeners were invited to choose only one of the two stimuli, after a complete listening of both, by pressing a button on the left (if poetry was detected in the first stimulus) or right (if poetry was detected in the second stimulus), to distinguish poetry between the two stimuli.

More in detail, the task included the following manipulated data: 5 audios of 1 poetess-Caproni-poetry (1 natural synthetic+4manipulations), 5 audios of 1 poetess-Caproni-prose (1 natural synthetic + 4manipulations), 5 audios of 1 announcer-Caproni-poetry (1 natural synthetic + 4manipulations), 5 audios of 1 announcer-Caproni-prose (1 natural synthetic + 4manipulations), 3 audio of 1 announcer-Ungaretti-poetry (1 natural synthetic + 2 manipulations), 3 audio of 1 announcer-Ungaretti-prose (1 natural synthetic + 2 manipulations), 3 audio of 1 actor-Ungaretti-poetry (1 natural synthetic + 2 manipulations), 3 audio of 1 actor-Ungaretti-prose (1 natural synthetic + 2 manipulations).

The task submission was carried out thanks to the involvement of 40 students of the Dep. of Foreign Languages and Literatures and Modern Cultures of the Univ. of Turin (General Linguistics Course, Master, Prof. A. Romano). They have been informed and consented to the study to be conducted, under a general University agreement and according to the GDPR.

After the training task (with double listening to each stimulus) each student participated in the discrimination task.

2.3. Results

The study results showed a major convergence in the recognition of poetry from prose[2]. In particular, with the discrimination task participants reached 72.6% of correct answers (23.55/32 correct answers). Limited are the cases of less than half of correct answers and all correct or wrong answers. On the other hand, the margin of chance (16/32) is almost absent.

The results highlighted a different reception according the reading styles of the three speakers: in particular, the reception of the poetess's reading appeared to be quite distinct from the actor's reading, which was the most frequently recognised, followed by that of the announcer-actress. Furthermore, the stimulus pair with the highest number of correct responses was that of the actor-reading of Ungaretti, with manipulation of duration (average). The percentages of recognised stimuli were 80% for the actor's voice, 76% for the announcer's voice, and 67% for the poetess' reading (Fig. 3).

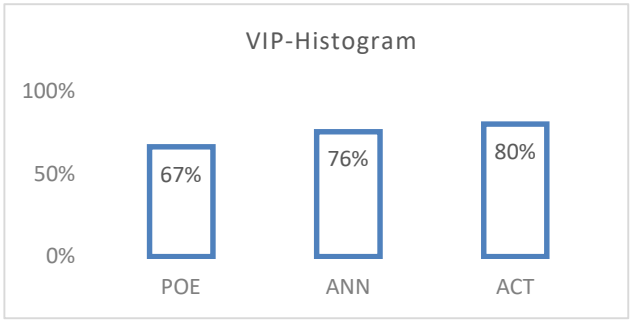

Figure 3: VIP-Histogram. Percentage of correct answers per vocal type.

Considering instead the percentages of correct responses concerning manipulation (see Fig. 4), we noted that the highest percentage of correct responses is found with the synthetic stimuli with flattening of f_0, followed by the unaltered stimuli and the stimuli with flattening of syllable mean durations (1st type).

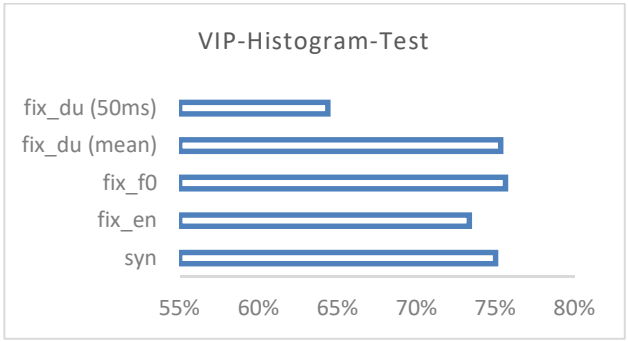

Figure 4: VIP-Histogram-test. Percentage of correct answers per type of stimulus.

Slightly lower are the results for the stimuli with energy flattening and even lower are those in which we adopted a second duration flattening (50 ms). In conclusion, speech rate, mean f_0, and intonation patterns appear to be central elements in recognition, even when syllable duration is normalised.

Two statistical ANOVA tests enabled verification of the significance of this analysis. In the first, the *type of speaker* factor (actor, announcer, poet) was fixed; in the second, the factor of the *type of manipulation* was fixed together with the *type of poem* read.

The first test showed that the type of voice had a statistically significant effect: the number of correct answers has been higher depending on the type of voice (higher with actor-voice, medium with announcer-actress, low with the poet): $F(2, 29) = 7.794$, $p<0.002$. We report in Fig. 4 the related Distribution Plot (ANOVA).

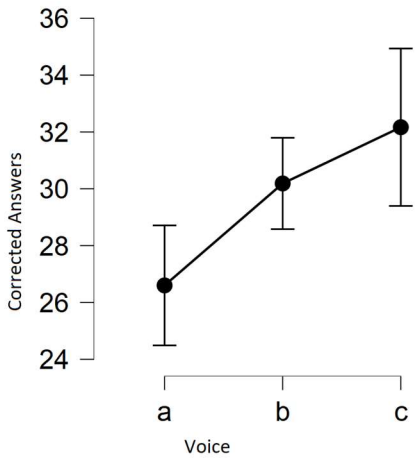

Figure 5: Distribution Plot (ANOVA) (a-poetess, b-announcer/actress, c-actor).

In Fig. 6 we report also the related Q-Q-plot, which shows the standardised residuals fit on the diagonal, suggesting that the hypotheses, normality, and linearity were not violated.

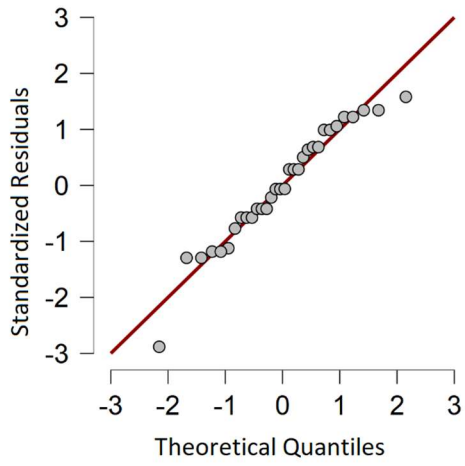

Figure 6: Q-Q-plot related to the first ANOVA test.

The second ANOVA test, with the fixed factors of manipulation (5 types set out above) and type of poem (Ungaretti/Caproni's compositions) shows that the type of manipulation test is statistically non-significant: F (4, 22) = 0.301, p<0.874. In contrast, the type of poem turns out to be a significant factor, F (1, 22) = 7.961, p<0.010. Finally, a non-significant interaction is given by the interaction of the two factors instead, F (4, 22) = 0.711, p<0.593. We report in Fig. 5 the related Distribution Plot (ANOVA).

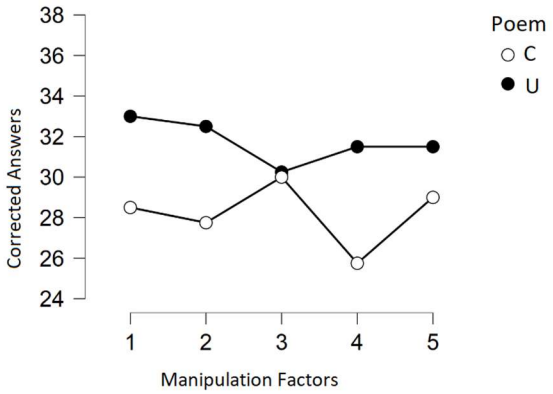

Figure 7: Second Distribution Plot (ANOVA) (5 Manipulation Factors).

This result confirms our hypothesis of an influence of the textual-formal and content difference (e.g. the substantial metrical difference between a text in hendecasyllables, such as Caproni's poem and one with various meters, as well as the clear linguistic difference) on different performances and receptions in the reader, conditioning the prosody (perceptible also in the delexicalised stimuli). At the same time, different voices (and ways to manipulate them) represent a further element of condition and influence in reception. In this way, the theme of the text and its form, together with the individual vocality, can be considered elements affecting the different modes of reading and their perception.

Furthermore, it can be deduced that excessive alteration of duration (too fast) prevents discrimination, as vocalic duration is a determining element and is considered to be distinctive. It also turns out that speech rate, mean f_0, and intonation patterns are central elements in recognition, even though the syllable duration is normalised.

3. CONCLUSION

This perceptual investigation showed that recognition of poetry from prose, through a metrical reading of the composition, is possible, even in the presence of delexicalised synthetic stimuli and possible further alterations of their acoustic parameters. This type of experimental task has proved to be a useful tool to deepen the acoustic dimension of the literary text.

It has emerged that the percentage of exact recognition of the poetry reading (with "metric" style) is greater where f_0 and duration are smoothed out, as well as in the unaltered synthetic stimuli.

The results highlighted that the perception of poetry prosody passes through recognisable and specific traits: first and foremost a specific speech rate, that appears central in poetic reading, which cannot be separated from a specific melodic use of the voice, melodically more homogenous, and a larger dimension of pauses.

Furthermore, the selected audience, familiar with literary and linguistic studies, easily associated poetry with a primarily actorial style. This result led to the hypothesis that a non-expert audience predominantly depicts its imagery of the poetic voice with this vocal modality, better known because it is usually proposed as a support in teaching. It must be connected to a historical tradition of actors' poetry reading characterised by a recitative-declamatory approach and a divergent tendency in poets' reading to increase the speech rate, in general terms.

The little association of the poetic voice with the voice of a poet(ess) can be understood as a manifestation of a lesser familiarity in listening to this kind of vocality and could be ascribable to the Italian didactic tradition mentioned above, on the whole, little linked to contemporary poets and their readings in the last twenty years. We, therefore, believe that a consolidated tradition of readings by authors and actors of the second half of the Twentieth century could be considered more settled in the ear of listeners. It would be related to stylistic features, as an emphatic and declamatory style, and slower speech rates (see [15]). In the common imagination, it would better correspond to the "painting" of the poetic voice, which would present a specific cadence in which the verse retains its distinguishable measure even when it is free and files are manipulated.

The perceptual dimension is to be understood as a significant part of what we consider a possible literary vocal imagery. A broadening of the research to a richer variety of texts and authors, as well as a larger number of vocal samples per category and speakers categories, together with a more differentiated listening audience and further differentiations of stimuli's combination, would represent a further step to describe the vocal literary imagery. As much as we believe that it can only be truly complete with the support of interdisciplinary studies, including cognitive sciences, we believe that research such as this can be considered an initial contribution to exploring the discrimination between poetry and prose (and their respective more typical reading styles).

4. REFERENCES

[1] Fónagy, I. 1987. Le lettere vive in poesia. In: *Le lettere vive*, 289-342. Bari: Dedalo.

[2] Bolinger, D. 1989. *Intonation and its uses.* London: Edward Arnold.

[3] Bröggelwirth, J. 2007. *Ein rhythmisch-prosodisches Modell lyrischen Sprechstils*. Ph.D. Dissertation, Universität Bonn. Then in: Bröggelwirth, J. 2010. *Ein rhythmisch-prosodisches Modell lyrischen Sprechstils.* Munich: GRIN.

[4] Wagner, P. 2012. Meter Specific Timing and Prominence in German Poetry and Prose. In: Niebuhr, O. (ed.) *Prosodies – Context, Function, Communication*, 219-236. Berlin/New York: De Gruyter.

[5] MacArthur, M., Miller, L. M. 2018. After Scansion: Visualizing, Deforming and Listening to Poetic Prosody. *Stanford ARCADE. Literature, the Humanities, & the World*, Dec. 13. https://arcade.stanford.edu/content/after-scansion-visualizing-deforming-and-listening-poetic-prosody.

[6] Hussein, H., Meyer-Sickendiek, B., Baumann, T. 2020. Free Verse and Beyond: How to Classify Post-modern Spoken Poetry. *Proc. 10th International Conference on Speech Prosody*. Tokyo (Japan), May 2020, 690-694.

[7] 't Hart, J., Collier, R., Cohen, A. 1990. *A Perceptual Study on Intonation. An Experimental Approach to Speech Melody*. Cambridge: Cambridge University Press.

[8] Romano, A. 1997. Persistence of prosodic features between dialectal and standard Italian utterances in six sub-varieties of a region of Southern Italy (Salento): first assessments of the results of a recognition test and an instrumental analysis. *Proc. 5th European Conference on Speech Communication and Technology (Eurospeech 1997)*, 175-178, doi: 10.21437/Eurospeech.1997-71

[9] Romano, A. 1999. Rising-Falling contours in Speech: a Metaphore of Tension-Resolution Schemes in European Musical Traditions? Evidence from Regional Varieties of Italian. *Atti del 8th workshop CSNLP – "Language, Vision & Music" (Galway, Irlanda, 8-12 Agosto 1999)*, 232-238.

[10] AMPER *Project-Atlas Multimédia Prosodique de l'Espace Roman* http://www.lfsag.unito.it/amper/amper.html

[11] Moulines, E., Werner, V. 1995. Time-domain and frequency-domain techniques for prosodic modification of speech. In: Bastiaan Kleijn, W., Paliwal, K. K. (eds.) *Speech coding and synthesis*, 519-555. Amsterdam: Elsevier.

[12] Van Bezooijen, R., Gooskens, Ch. 1999. Identification of language varieties: The contribution of different linguistic levels. *Journal of language and social psychology*, 18(1), 31-48.

[13] Knoll, M. A., Uther, M., Costall, A. 2009. "Effects of low-pass filtering on the judgment of vocal affect in speech directed to infants, adults and foreigners". *Speech Communication*, 51(3), 210-216.

[14] Colonna, V. 2021. *Voices of Italian Poets. Analisi fonetica e storia della lettura della poesia italiana dagli anni Sessanta a oggi*. PhD Dissertation, University of Genoa-University of Turin.

[15] Colonna, V. 2022. *Voices of Italian Poets. Storia e analisi fonetica della lettura della poesia italiana del Novecento*. Alessandria: Dell'Orso.

[16] Bologna, C. 1990. Les interprètes. *Le grand atlas des littératures a été réalisé par Encyclopoedia Universalis*, 80-83. Paris: Encyclopaedia Universalis.

[17] AMPER Routines: http://www.lfsag.unito.it/amper/amper.html.

[1] This study takes up and reworks part of the first author's PhD Dissertation.

[2] The training test gave the same result: participants reached 68% of correct answers related to the identification of poetry and 66% related to prose.