

POEM DECLAMATION IN TWO VARIETIES OF PORTUGUESE: THE LINK BETWEEN PROSODY AND PLEASANTNESS

Plinio A. Barbosa

Dep. of Linguistics, Instituto de Estudos da Linguagem, University of Campinas, Brazil
pbarbosa@unicamp.br

ABSTRACT

In poem declamation, the appropriate use of prosody is essential to cause pleasure. The declamation of three poems in two varieties of Portuguese by ten Brazilian and ten Portuguese speakers balanced for gender is used as a corpus for evaluating the degree of pleasantness by listeners from the same language variety. Six prosodic acoustic parameters related mostly to melody and timing varied largely across the subjects, which is the main source of variability. The evaluation of degree of pleasantness revealed that pause-related and voice quality parameters are the best predictors of pleasantness, but the two varieties differ in the specific parameters and direction of change that increases pleasantness.

Keywords: Poem declamation, Prosody, Portuguese, Pleasantness

1. INTRODUCTION

According to [1], “the evaluation or production of beauty, ugliness, harmony, elegance, fitness or charm is governed by a set of factors such as stimulus symmetry, complexity, novelty, familiarity, artistic style, social status and individual preferences.” Elements such as novelty and complexity are associated with prosodic aspects such as the rhythmic structure and the intonation organisation of recited verses, which have effects of surprise and expectation, among others. These effects are coupled together in such a way as to make a performance, such as a recitation, sounds pleasant. Several prosodic parameters seem to have a role in the positive appreciation of a voice. In German, for instance, [2] found a relationship between soft, breathy speech and a lower spectral center of gravity with the growth of positive appreciation by listeners. Also in German, the acoustic measures by [3] seem to confirm these very findings by showing that voices with lower harmonic-to-noise ratios and higher jitter, both related to breathiness, got more pleasant appreciations.

In the specific case of poetry declamation, [4]

showed that the realisation of pauses in places not predictable from syntax increased judgments of pleasure, whereas [5] proposed that what characterises the acoustics of poetry declamation are slower speech rates, more frequent pauses, short tone units, relatively equal-length units, low average pitch, and narrow pitch range.

As regards work on European Portuguese (EP), [6] built a Machine Learning system for classifying and predicting the degree of pleasantness of 3-min audio excerpts from prosodic- and voice-quality acoustic parameters. Their results showed that jitter, shimmer, F0 mean, slope of intensity and maximum of F0 slope were the best predictors of pleasantness.

By investigating poetry declamation in Brazilian Portuguese (BP), [7] evaluated the perceptual effects of the production choices of two professional speakers (male and female) reciting the poem “Soneto da Fidelidade” (Sonnet of Fidelity) by Vinícius de Moraes. The acoustic analysis of both readings revealed that whispery voice, F0 narrow range and a great number of silent pauses were the production strategies used by the female speaker to express grief and sadness, whereas the male speaker employed much varied intonation patterns to express liveliness. In a follow-up study of the effects of the recitation of the same poem [8] evaluated by eight non-professional speakers, pleasantness was moderately relevant as a subjective parameter related to voice quality and the slope of LTAS.

Working on the acoustic of poetry declamation in both BP and EP, [9] showed that the degree of pleasantness attributed by Brazilian listeners to the recitations in BP was explained by three parameters in order of importance: rate of pause (the lower, the more pleasant), slope of LTAS (the faster the slope, the more pleasant) and pause duration (the longer, the more pleasant) with an additional minor prediction power for F0 median in the case of male reciters. On the other hand, the appreciation of EP recitations by Portuguese listeners depended only on female reciters’ pause duration, whereas males’ recitations were judged more pleasant due to variation in spectral emphasis, F0 median (the lower, the more pleasant in both cases), and pause

duration (the longer, the more pleasant), in that order of importance.

Based on the reviewed literature and the intuition that poem declamation is a genre in its own that depends less on the content for its expression, our main hypotheses are: H1: There is a cross-variety difference in terms of prosodic-acoustic parameters that explain pleasantness in BP and EP; H2: Pause-related and voice quality parameters are the main parameters for explaining the degree of pleasantness attributed by the listeners, regardless of the theme or the poet; H3: The gender of the reciter affects how pleasantness is evaluated; H4: Rate of pause in BP recitations, and pause duration and spectral emphasis in EP recitations are the most important predictors of pleasantness. The main research question is: What are the acoustic correlates of pleasantness in poetry declamation in BP and EP?

2. METHODOLOGY

The PROS-POIESIS corpus is formed by the declamations of four poems: “Quando vier a primavera” (When Spring comes), with 155 words, and “O amor é uma companhia” (Love is a company), with 111 words, both by Alberto Caieiro (one of Fernando Pessoa’s heteronyms, a Portuguese poet), as well as two poems by Adélia Prado, a Brazilian poetess: “Momento” (Moment), with 108 words, and “Tempo” (Time), with 71 words. The four poems were recited by ten Brazilian speakers and ten Portuguese speakers in balanced gender. The set of reciters of the three latter poems was formed by exactly the same persons (called here the Current Set), but this is not the case for the reciters of the first poem, due to the unavailability of the same persons in the moment of the second set of recordings.

Due to the reference of death in “Quando vier a primavera”, its valence was labeled as negative, in contrast to the positive one of the second poem “O amor é uma companhia”, whose theme is the poet’s joy for loving someone. The two poems by Adélia Prado were chosen in order to contrast the positive valence of “Tempo” and the negative valence of “Momento”. Not a single reciter has voice professional training and all of them have between 25 and 50 years of age to avoid an age group with effects of vocal aging [10] that usually affect melodic, voice quality and temporal parameters such as speech and pause rate.

Due to the Covid-19 pandemic, the participants themselves used the Easy Voice app in their own cell phones to make all recordings. Because this app allows choosing among different codifications,

instructions were given to record all audio files in PCM format (WAV) at a sampling rate of 48 kHz. The author, who is a trained phonetician, further evaluated all audio files. The recordings were resampled at 16 kHz and levelled to the same maximum intensity level at 65 dB.

2.1. Acoustic parameters

The Prosody Descriptor Extractor script for Praat [11] developed by the author [12] was used to extract 22 prosodic-acoustic parameters from a verse-based segmentation of the poems where silent pauses were also segmented. The declamation of each one of the three poems of the Current Set lasted from 31 to 61 s.

For each verse, the prosodic parameters extracted were the following: 12 descriptors of F0: median, semi-amplitude between quartiles (F0SAQ), minimum and maximum, standard-deviations of values and time of F0 local peaks, mean peak rate and mean peak bandwidth, mean and standard-deviations of F0 rates of rises and falls; two intensity descriptors: spectral emphasis [13], and coefficient of variation of total intensity; four voice quality descriptors: Harmonic-to-Noise ratio, HNR (a correlate of breathiness), long-term averaged spectrum slope, computed by the difference in mean energy between the bands 0-1 kHz and 1-4 kHz, jitter and shimmer; and, finally, four temporal descriptors: pause duration and pause rate, speech and articulation rates. Pause rate is studied here by inter-pause interval (IPI) values, because pause rate is the inverse of IPI mean.

2.2. Perception tasks

Two perception tests were carried out: one Likert-scale test for evaluating the degree of pleasantness applied to ten Brazilian listeners and containing declamations by Brazilian reciters, and one Likert-scale test for evaluating the degree of pleasantness applied to ten Portuguese listeners and containing declamations by Portuguese reciters. For that task, each participant listened to the entire declamation of all reciters of s/he variety, which gave 30 declamations to be evaluated per variety (10 reciters x 3 poems). The scale for pleasantness evaluation varied in five degrees from “very unpleasant” (“muito desagradável” in Portuguese, degree 1) to “very pleasant” (“muito agradável” in Portuguese, degree 5) with the neutral response having degree 3. For the perceptual tests, the listeners were selected in the same age range as the reciters to avoid effects of differences in evaluation from younger or older

age groups. In the present study, only the recitations and the evaluations by the listeners of the Current Set are considered for analysis.

2.3. Statistical analyses

To test our hypotheses, three kinds of statistical models were run in the R (R Development Core Team) software. Due to the non-normality of the parameter's distributions, a non-parametric 2-Way ANOVA test, the Scheirer-Ray-Hare (SRH) test, was used for evaluating the mean differences according to two factors: Poet (Adélia Prado vs. Alberto Caieiro) and Valence (positive vs. negative). The models were built on the subsets split according to gender and variety, in each case.

Mixed-effect models, also split by gender and variety, were built in order to evaluate the strength of the two fixed-factors, Poet and Valence, and the random factor, Subject for explaining each parameter's variance. In each case, the rank of the parameter value was used instead, due to the non-normality of the distributions.

The third set of statistical models was based on logistic regression. Two final logistic regression models, one for each variety, corresponding respectively to the two Likert-scale tests presented in the previous section, were retained that predicted the degree of pleasantness from the significant predictor variables (the prosodic parameters). For doing so, the degrees from 1 to 5 were respectively transformed to 0 to 100% in intervals of 5%, allowing the use of a logistic model. Nagelkerke pseudo-correlation measures were used to evaluate the degree of explained variance of these models, which are measures of effect size. In all tests, a 0.05 level of significance was used for decision as regards statistical differences.

3. RESULTS

3.1. Production

The SRH models revealed some significant differences depending on valence and/or poet. Tables 1 and 2 show the significant parameters common to EP and BP: F0 maximum (in Hertz), F0 rise rate (Hertz/frame), F0 fall rate (Hertz/frame), articulation rate (syll./s), pause duration (ms), and IPI (s). When the difference refers to one gender in particular, it is indicated between parentheses (M or F). When the difference refers to valence this is indicated by *n* from negative, or *p* from positive. The same is done in the case of differences related to poet (CA = Alberto Caieiro, AP = Adélia Prado).

The mean values are also given.

parameter	BP	
	contrast	p
F0max	n 150 / p 163	0.008 (M)
F0 rise rate	CA 2.2 / AP 2.5	0.0008 (F)
F0 fall rate	CA 3.2 / AP 4.2	0.005 (F)
	n 1.6 / p 2.1	0.003 (M)
art. rate	AP.p 5.3 / Others 5.8	0.001
pause dur.	n 476 / p 617	0.05 (M)
IPI	AP.n 2.2 / Others 2.6	0.02 (F)

Table 1: Significantly different parameters for valence and/or poet in common with EP.

Besides the parameters differing in both varieties, there are significant parameters exclusive to a particular variety. In BP, the standard-deviations of F0 peak values ($p = 0.009$) and temporal positions ($p = 0.001$) differ for Valence in the case of female reciters, whereas speech rate is lower for the positive poem by AP (4.2 syll./s vs. 4.5 syll./s for the other poems, $p = 0.01$) in both genders. In EP, F0 mean peak width is higher for the declamations of AP ($p < 0.03$) for male reciters only.

The mixed-effect models revealed that both Poet and Valence and their interaction explain a maximum of only 4% of variance (the highest values are for IPI and pause duration in both BP and EP), considering all significant parameters, whereas the random factor (Subject) explains between 15 and 85% of the total variance, indicating that the subject is the main source of variability. The parameters whose variance is better explained by Subject are melodic (F0 median: around 80%) and voice quality-related (spectral emphasis: 75% in BP and LTAS slope: 72% in EP). Subjects explain between 20 and 30% of the variance of temporal parameters such as speech rate and articulation rate, whereas those of pause duration and IPI relate between 50 (BP, IPI) and 85% (EP, pause duration) with the subjects.

parameter	EP	
	contrast	p
F0max	n 222 / p 231	0.07 (F)
F0 rise rate	CA 2.3 / AP 2.9	0.02 (M)
F0 fall rate	CA 1.4 / AP 2.2	0.0002 (M)
art. rate	AP 6.0 / CA 6.6	10^{-5}
pause dur.	AP 498 / CA 641	0.005 (M)
IPI	n 1.9 / p 2.2	0.003

Table 2: Significantly different parameters for valence or poet in common with BP.

3.2. Perception

The best logistic regression models to predict pleasantness from the acoustic parameters differ across language variety. For BP, in order of importance: IPI (the longer, the more pleasant); spectral emphasis (the lower, the more pleasant); HNR (the higher, the more pleasant); shimmer (the lower, the more pleasant); intensity c.v. (the lower, the more pleasant); and jitter (the higher, the more pleasant). The Nagelkerke pseudo-correlation for this model was 78%. This model does not differ significantly between male and female reciters.

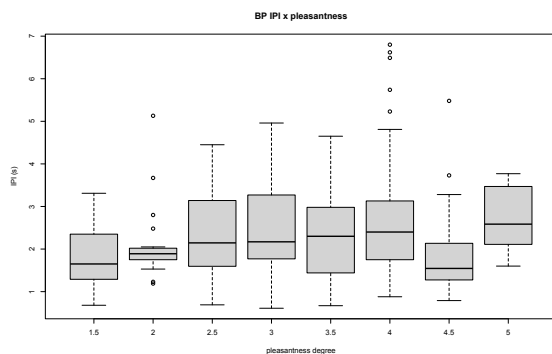


Figure 1: Boxplots of IPI vs. pleasantness in BP.

To illustrate the importance of IPI for pleasantness in BP, see Figure 1. Observe the increase in IPI median for higher values of IPI, with the exception of degree 4.5, which is exclusive for female reciter F2BP evaluations. In the repository given in this paper the reader can appreciate the lowest F0 variability among the female reciters, a feature that could explain her higher degree of pleasantness.

For EP, the models depend on the gender of the reciter. In the case of females, the significant parameters in order of importance are: pause duration (the longer, the more pleasant); rate of pause (the higher, the more pleasant); shimmer (the higher, the more pleasant); LTAS slope (the lower, the more pleasant, which means, less energy in the higher frequencies). For males, only pause duration (the longer, the more pleasant), and LTAS slope (the higher, the more pleasant, which means, higher energy in the higher frequencies) were significant. The Nagelkerke pseudo-correlations for both models were circa 90%.

4. DISCUSSION

The valence of the poem has a relation to the prosodic parameters. If you look at Tables 1 and

2, F0 maximum (BP males and EP females) and F0 falls (BP males) are higher in positive poems. This could be related to the production of more lively recitations in these poems. Lower IPIs are found in negative poems in at least one poem in both varieties as well. A relation with content cannot be discarded, though, because the recitations in both varieties have higher F0 rises and falls and lower articulation rates for the poems by Adélia Prado.

The findings by [2] and [3] for German are partially the same for Portuguese: in BP, recitations with lower spectral emphasis (less effortful voices) and higher jitter are also preferred. As for HNR, a higher HNR in BP is preferred, though. The role of both jitter and shimmer is confirmed for BP and EP females, as in [6]. Because voice quality parameters were found to be significant for explaining pleasantness in BP, this confirms the findings by [8].

The poetic formula by [5], which proposes that slower speech rates, more frequent pauses, low average pitch, and narrow pitch range, inter alia, are characteristic of poetry, was only related to the fact that more frequent pauses in EP female recitations were more pleasant. The valence and different contents of the recitations presented here do not vary in terms of these parameters as well, which probably could signal that they are not that relevant for present-day poetry declamations.

As regards our hypotheses, H1 to H3 are confirmed and H4 is partially confirmed: although the rate of pause in BP recitations and of pause duration in EP are the most important predictors of pleasantness, spectral emphasis is not a significant predictor in EP but still the second best predictor of pleasantness in BP.

Our findings can contribute as guidelines for coaching aiming at instructing actors, entrepreneurs to have more pleasant voices for the sake of entertaining or persuasion.

5. ACKNOWLEDGMENTS

All stimuli are publicly available for listening at: figshare.com/authors/Plinio_Barbosa/11320902. This research was funded by the CNPq, grant # 302194/2019-3. The study was approved by the Comitê de Ética em Pesquisa of the University of Campinas (protocol # 61430122.9.0000.8142). We thank Ana Rita Valente, Quintino Lopes and Catarina Oliveira in Portugal and Sandra Madureira in Brazil to help with participant recruitment. I also thank Alicia Crochiquia for preparing the online links for the perception tests.

6. REFERENCES

- [1] T. Jacobsen, “Beauty and the brain: culture, history and individual differences in aesthetic appreciation,” *Journal of Anatomy*, vol. 216, no. 2, pp. 184–191, 2010.
- [2] B. Weiss and F. Burkhardt, “Voice attributes affecting likability perception,” in *Proceedings of INTERSPEECH 2010*, Chiba, Japan, 2010, pp. 2014–2017.
- [3] J. M. Carlsen, M. Franke, L. M. Huttner, and A. Radtke, “How to measure a pleasant voice,” in *Proc. of the Conference Phonetics and Phonology in the German Language Area (P&P14)*, Vienna, 2018, pp. 22–26.
- [4] S. Postarnak, S. L. Castro, and S. Silva, “The role of linguistic prosody in the responses to recited poetry,” *Scientific Study of Literature*, vol. 10, no. 2, pp. 228–249, 2020.
- [5] P. B. Byers, “A formula for poetic intonation,” *Poetics*, vol. 8, pp. 367–380, 1979.
- [6] L. Pinto-Coelho, D. Braga, M. Sales-Dias, and C. Garcia-Mateo, “On the development of an automatic voice pleasantness classification and intensity estimation system,” *Computer Speech & Language*, vol. 27, no. 1, pp. 75–88, 2013.
- [7] S. Madureira, “Reciting a sonnet: Production strategies and perceptual effects,” in *Proc. Speech Prosody Conference 2008*, Campinas, Brazil, 2008, pp. 697–700.
- [8] P. Menegon, S. Madureira, and M. Fontes, “Análise de interpretações orais de um poema quanto aos aspectos expressivos,” *Intercâmbio*, vol. 49, pp. 113–137, 2021.
- [9] P. A. Barbosa, “Pleasantness and wellbeing in poem declamation in European and Brazilian Portuguese depends mostly on pausing and voice quality,” *Front. Commun*, 2022, 7: 855177. DOI: 10.3389/fcomm.2022.855177.
- [10] E. T. Stathopoulos, J. E. Huber, and J. E. Sussman, “Changes in acoustic characteristics of the voice across the life span: Measures from individuals 4 - 93 years of age,” *Journal of Speech, Language, and Hearing Research*, vol. 54, no. 4, pp. 1011–1021, 2011.
- [11] P. Boersma and D. Weenink. (2021) Praat: doing phonetics by computer. [Online]. Available: <http://www.praat.org/>
- [12] P. A. Barbosa. (2020) Prosody Descriptor Extractor. [Online]. Available: <https://github.com/pabarbosa/prosody-scripts/tree/master/ProsodyDescriptorExtractor/>
- [13] H. Traunmüller and A. Eriksson, “Acoustic effects of variation in vocal effort by men, women, and children,” vol. 107, pp. 3438–3451, 2000.