The Acoustics of Pleasantness in Poetry Declamation in Two Varieties of Portuguese

Plínio A. Barbosa

Institute for Language Studies, University of Campinas, Brazil
pabarbosa.unicampbr@gmail.com

Abstract

This work investigates the link between a poem declamation and the sensation of pleasantness in two varieties of Portuguese: European and Brazilian Portuguese. A group of ten Brazilian and ten Portuguese reciters had their performances evaluated by a group of ten Brazilian and ten Portuguese listeners invited to judge the degree of pleasantness. The perceptual evaluation was done in two ways: first, the listeners of the respective linguistic community carried out a Likert-scale test to attribute a degree of “very unpleasant” to “very pleasant” in five steps for the declamations of Alberto Caeiro’s poem “Quando vier a primavera”. Then, the listeners of the two varieties together carried out a discriminant test to point which one of two declamations of the same chunk from the poem was more pleasant. Both reciters and listeners were balanced in gender. A set of 22 prosodic-acoustic parameters was extracted from the verses and taken as predictors in logistic regression and Linear Discriminant Analysis models for explaining the perceptual evaluations. Results point to pause duration and rate, articulation rate, rate of F0 falls, spectral emphasis and LTAS slope as the most relevant predictors of pleasantness. Implications for an acoustic of pleasantness are discussed.

Index Terms: Vocal aesthetics, Prosody, Pleasantness

1. Introduction

According to [1], speaking style can be defined as a “differentiation in the way of speaking”, a proposal to which we must add that it is also necessary that it constitute a genre with common characteristics, so that one can speak of the broad styles of reading, narration, religious sermon, political discourse, among many others. The vocal difference between speaking styles is related to changes in voice quality, speech rhythm and intonation, according to studies carried out in the field for decades (cf. [2] for a review). Among these, professional styles are identified, such as TV and radio broadcasting, sports broadcasting, acting [3], to which we must include the declamations of the verses, an activity intended to provoke in the listener effects of pleasantness, a result also confirmed by [8], which adds that women also prefer men who have higher speech rates (see also the same result in [9, 10]). On the other hand, [7] reported that men find the speech of women with higher mean F0 more attractive. Another relation to gender is the fact that [11] found that women generally found the speech of other women more pleasant, although this result contrasts with previous research [12].

Work on German showed that attractiveness, regardless of gender, seems to increase with the variability of F0 [13], as well as with the absence of disfluencies [14, 9]. In a previous work, [15] found a relationship between soft, breathy speech and a lower spectral center of gravity with the growth of positive appreciation by listeners.

As can be seen from this overview, there are issues related to the link between prosodic-acoustic parameters and the sensations of pleasantness and attractiveness in speech that seem to be recurrent, at least in English and German, while others need further study such as the question of appreciation by men and women of the speech of people of the same gender or not. Despite recent studies on voice quality parameters which showed that lower harmonic-to-noise ratio and higher jitter in German are related to an increase in pleasantness [16], additional acoustic parameters could explain aesthetic appreciation such as voice quality measures related to vocal effort, pausing, intensity, among others. Some of these measures are examined here.

Besides extending the spectrum of potential prosodic parameters of pleasant speech, research on Romance languages such as Portuguese could help contribute to increase knowledge on the acoustics of pleasantness. That it why the present work investigates the link between the acoustics and the perception of pleasantness of poem declamation in two varieties of Portuguese. Pointing out acoustic characteristics of the declamation that contribute to an increase in pleasantness could provide elements for the training of professional speakers.

2. Goals and hypotheses

The main goal of this work is to investigate the link between the declamation of a poem from Portuguese writer Alberto Caeiro and the sensation of pleasantness in both European (EP) and Brazilian Portuguese (BP). As specific goals, it intends to evaluate (1) how the declamation-pleasantness link works in EP and BP, (2) which prosodic-parameters contribute to an increase of pleasantness in the two varieties of Portuguese according to male and female listeners, (3) possible gender differences of speakers and listeners, (4) whether an actor from each culture has his declamation considered more pleasant than those of lay reciters, (5) whether pleasantness comparison across verses’
recitation reflect pleasantness degree evaluation of the whole poem declamation. Based on the studies reported in the previous section, the hypotheses of this work are:

1. Women find the declamation of men with lower mean F0 more pleasant, while men find it more pleasant that of women with higher mean F0;

2. Although the literature mentions that the higher speech rate, the more pleasant speech is, we hypothesise the opposite in the case of a poem declamation. In fact, slowing down the recitation is related to an increase in care with what is being said, thus contributing with a positive evaluation. Then, the longer the average duration of pauses, the more pleasant speech is;

3. Portuguese listeners appreciate the declamation by Brazilians more than Brazilians appreciate the declamation by Portuguese speakers, due to the greater exposure in Portugal to the speech of Brazilians on Portuguese TV, especially Brazilian soap-operas;

4. The breathier the speech, the more pleasant the recitation is;

5. The more variable F0, the more pleasant the recitation is;

6. The softer, with less vocal effort the speech, the more pleasant the recitation is;

7. The acoustic parameters related to pleasantness are the same and in the same degree of relative importance in the two varieties of Portuguese, in the absence of evidence to the contrary;

8. Men and women use the same prosodic-acoustic parameters to make speech better appreciated by the listener;

9. Actors’ voices have more pleasant voices than those of lay reciters.

3. Methodology
The PROS-POIESIS corpus is formed by the declamation of the poem “Quando vier a primavera” (When Spring comes), by Alberto Caetano, one of Fernando Pessoa’s heteronyms, by ten Brazilian speakers and ten Portuguese speakers in balanced gender. All speakers do not have professional training for voice. Due to the pandemic, all recordings were made by the participants themselves using a cell phone with good PCM-coded recording quality which was evaluated by the author, a trained phonetician. The recordings were resampled at 16 kHz and levelled to the same maximum intensity level at 65 dB. As control recordings, the declamations of the same poem by Brazilian actor Ivan Lima and by Portuguese actor Pedro Lamas are included in the dataset. These recordings were downloaded from the Internet and are available publicly. It is expected that they get the highest scores of pleasantness due to their professional voices.

3.1. Acoustic parameters
The Prosody Descriptor Extractor script for Praat [17] developed by the author [18] was used to extract 22 prosodic-acoustic parameters from 14 chunks from one to two verses of the spoken poem including the silent pause at the end, when applicable. The segmentation of the chunks was the same across reciters and the reading of the poem lasted from 60 to 95 s. For each chunk there were 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, a correlate of vocal effort [19] 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 (LTAS-slope, computed by the difference in mean energy between the bands 0-1 kHz and 1 kHz-4 kHz), jitter and shimmer, and finally four temporal descriptors: silent average duration and pausing rate, speech and articulation rates.

3.2. Pleasantness evaluation
Two perception tests were carried out: a differential semantic scale test for the entire reading of the poem evaluated by the respective linguistic community and a discrimination test for the passages corresponding to the verses with all listeners from the two linguistic communities.

As part of the semantic scale test, we included the declamation of the Brazilian actor to the declamations from ten lay Brazilian speakers which were evaluated by ten Brazilian listeners balanced in gender in a five-point Likert scale from “very unpleasant” (degree 1) to “very pleasant” (degree 5) passing by a neutral response (degree 3). Similarly, we included the declamation of the Portuguese actor along with the declamations from ten lay Portuguese speakers which were evaluated by ten Portuguese listeners also balanced in gender.

As for the discrimination test, a set of 44 pairs extracted from the recitations combining gender and speaker separated by a short musical tone was prepared. The chunks were extracted from the reading of the whole poem by the same speaker. The pair being compared was formed by the same chunk and language variety for avoiding a choice biased towards differences in content or variety. To evaluate the behaviour of listeners according to variety, the group of 20 listeners, from Brazil and Portugal, evaluated the same set. Their task was to say which of the verse readings is the one that sounds more pleasant, the first or the second.

For the two types of test, speakers and listeners were selected in the range between 25 and 50 years old to avoid the age group with effects of vocal aging [20]. Besides evaluating responses related to the gender of the listeners and reciters in the two tests, the discrimination test also allowed us to evaluate the possible effect of the variety used by the listener in evaluating pairs of verses recited in BP or in EP, whether the actors sound more pleasant, as well as whether the mean degrees attributed in the Likert scale test could predict the performance of the listeners in the discrimination test. Each listener participated online in the two tests mounted in the Survey Monkey platform with the help of a research assistant. The discrimination test was performed after the Likert scale test and the two tests lasted from 20 to 40 minutes.

3.3. Statistical models
To test our hypotheses, three kinds of statistical tests were run in the R project software [21]. Correlation tests were applied to verify whether Brazilian and Portuguese listeners answered in a similar way to the stimuli of the discrimination test, and whether these responses are correlated with the differences in degree from the previous test (more details in the next section).

Logistic regression models were built to predict degree of pleasantness from the prosodic-acoustic parameters, and Linear Discriminant Analysis (LDA) models were built to predict one of the two possible responses from the differences of the mean
values of each parameter from the two readings of the verse. In the two kinds of models the 5% level of significance was chosen. As for the models’ assumptions, the quasi-binomial family was used for prediction in the logistic regression models when there were important differences between degrees of freedom and amount of deviance. When this was not the case, the binomial family was used. Pseudo-correlation measures were used to evaluate the degree of explained variance. As for the LDA models, when two predictors correlate for more than 50%, just one of them was retained in the model.

4. Results

Responses of Brazilian and Portuguese listeners for the discrimination tests had correlations superior to 75% in all cases: the overall correlation was 78%; when the variety of the reciter was BP their correlation was 82%, and 76% when the variety of the reciter was EP. When both recitations of a chunk were from male speakers, the two groups of listeners correlated 78%, in comparison with 83% for female reciters and 76% for mixed-gender reciters.

The statistical results are presented after the results for the Likert-scale test. In this test, no differences were observed between male and female listeners when evaluating either male or female reciters as can be seen by checking the closeness of median responses in Figures 1 and 2.

If the direction of the differences in degree of pleasantness attributed to the whole poem’s reciter by the respective linguistic community is used as a prediction of the response in the discrimination test, the correlations between predicted-from-Likert-test and actual discrimination tests are 54% for Brazilian listeners and 63% for Portuguese listeners. One of the vectors was formed by the difference of the median degrees got by each reciter for the chunk extracted from the whole reading and, the other vector, the responses of the discrimination test (2, if the second reading is more pleasant than the first and 1, in the opposite case).

The logistic models considered several combinations of the 22 prosodic-acoustic parameters. Only the combination with a set of significant predictors explaining the maximum of variance when compared to all the other combinations was retained as the final model. Although, as just mentioned, the variety of the listener did not interfere in a relevant way with the results, the gender of the reciter in each spoken variety produced different models, that is why the results are presented genderwise for each group of reciters (Brazilian vs. Portuguese). The actors in each community got the highest median degree (5) for pleasantness.

The degree of pleasantness attributed to Brazilian female reciters by the Brazilian listeners is 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). Pseudo-correlation explains 42% of variance. The degree of pleasantness attributed to Brazilian male reciters by the same listeners is explained by four parameters: the same as for female reciters in the same order of importance plus F0 median in the last position (the higher, the more pleasant). The parameters explain 23% of variance.

As for the degree of pleasantness attributed to Portuguese female reciters by the Portuguese listeners, only pause duration was significant (the longer, the more pleasant), explaining 44% of the variance. As for results for male reciters given by the same listeners, three parameters in order of importance explain the data: spectral emphasis and F0 median (the lower, the more pleasant), and pause duration (the longer, the more pleasant), which explains 38% of variance.

As for the discrimination results, the overall prediction of the listeners’ choices is explained by four parameters in order of importance: articulation rate (faster rates are preferred), F0 falls (smoother slopes are preferred), spectral emphasis (lower values, softer readings are preferred), and with a minor load, slope of LTAS (sharper slope are preferred, which is related to less effort). Prediction for first or second preferred verse declamation got the same hits’ proportion: 75%. If the responses are separated according to the variety of the reciter, only the first three parameters of the overall model are significant in different orders of importance: the same of the overall model when Brazilian reciters were evaluated, with 88% of hits in both directions, and the following order when Portuguese reciters were evaluated: spectral emphasis (softer readings are preferred), F0 falls (smoother slopes are preferred), and articulation rate (faster rates are preferred). The preference for the first reading was predicted with 78% of hits and 75% of hits for the second reading choice. The median response by both Portuguese and Brazilian listeners for the chunk read by the Brazilian actor was for his reading in 6 out of 6 comparisons with the lay reciters,
and the median response by the two groups of listeners for the chunk read by the Portuguese actor was for his reading in 5 out of 6 comparisons with the lay reciters. A female lay (Portuguese) speaker was preferred over the Portuguese actor in this single exception.

Because measures of articulation rate and vocal effort, like spectral emphasis and LTAS slope, are more known in the literature, Figure 3 presents the F0 contours for the reading of a one-verse chunk by two Brazilian female speakers. The reading corresponding to the contour with sharper decreasing slopes from the F0 peaks was considered less pleasant. The direction of this choice happened 61% of all choices.

5. Discussion

The most relevant parameters for both the evaluation of the whole poem and the comparison of the respective chunk readings point to the classes of pause and tempo (pause duration, pausing rate and articulation rate), voice quality (spectral emphasis and LTAS slope) and melody (F0 falls), with a minor prediction power for F0 median in Brazilian and Portuguese male reciters. The degree of pleasantness attributed to the recitation of the whole poem by Portuguese female speakers seem to depend only on pause duration, though. Actors were preferred in all cases with a single exception, both in direct comparisons and in the whole poem recitation where pause rate and duration are relevant: the actors exhibit the longest median pauses. The Portuguese actor also has the lowest rate of pausing, the highest articulation rate and the smoothest F0 mean falls.

As for our hypotheses, Hypothesis 1 was not entirely confirmed because F0 median was significant only in the case of male speakers with higher pitch which were preferred by the two Brazilian genders of listeners, whereas male speakers with low pitches were preferred by the two Portuguese genders of listeners. Because speech rate was not significant and articulation rate predicted faster rate as being more pleasant, our hypothesis 2 was not confirmed. The high correlations of responses between Brazilian and Portuguese listeners broke down according to the variety of the reciter disconfirms hypothesis 3 as well. Hypotheses 4 and 5 were also not confirmed because HNR and F0SAQ were respectively not significant predictors. Hypothesis 6 is confirmed by the discriminant tests, because spectral emphasis predicted listeners’ preferences: the softer the declamation, the more pleasant it is.

As for Hypothesis 7, although there is a close correspondence between the acoustics of the reciters’ varieties of Portuguese, meaning that rhythmic parameters such as pause duration and articulation rate are significant predictors in the majority of models, with F0 median added in the case of male reciters in both varieties and measures of vocal effort, namely spectral emphasis and LTAS slope, the order of importance and specific combination of predictors is not the same in the two varieties of Portuguese. And because the predictors change depending on the gender of the reciter, Hypothesis 8 is not confirmed. Hypothesis 9 is confirmed because the actors’ declamation was preferred in all cases with a single exception when the Portuguese actor chunk is compared with the same chunk by a Portuguese female reciter.

As for the effect of the theme of the poem, namely the insignificance of the poet in comparison with the passage of the seasons, especially the return of Spring, although it does not favor a lively recitation, liveliness was present in several readings by our reciters. Due to that, and the high degrees of pleasantness and well-being that the majority of recitations received, we think that the content of the poem was not a determinant factor for explaining perception.

One key component of the acoustics of pleasantness seems to be rhythm of declamation because recitations with longer pause durations and lower pausing rates are more pleasant when our participants listen to the whole poem declamation. In contrast, in pair comparisons, higher articulation rates got more pleasant evaluation, which in accordance with the higher attractiveness of faster male speech judged by women. Our results show that, at least for pleasantness, higher articulation rates are preferred by both male and female listeners irrespective of the gender of the reciter.

6. Conclusions

The acoustics of pleasantness has some points in common with the acoustics of attractiveness when our results are compared with the findings reviewed in the Introduction mainly for English and German. One of these points is the role of tempo and vocal effort, as highlighted in the previous section. Despite minor differences in the evaluation of pleasantness depending on the variety of the reciter, the overall picture is quite the same, with a main role for tempo and voice quality through vocal effort acoustic correlates. The role of F0 falls in pair comparisons must be further investigated, but could also be related to slowing down, because smoother (slower) falls were considered more pleasant.

Because of some commonalities for both Germanic and Romance languages, it is possible that, at least in terms of Occidental culture, there are some common grounds in terms of the prosodic organisation of poem declamations to sound more pleasant.

7. Acknowledgements

I thank our reciters and listeners from the two sides of the Atlantic Ocean and the CNPq grant # 302194/2019-3. I am also in debt with research assistant Alice Crochiquia for her help with the online perception tests. This work is registered in the Ethics Committee under number 41315420.8.0000.8142.

8. References


