v. 4, n. 3, p. 44-53, 2015
ISSN 2237-9223

REVISTA BRASILEIRA DE CRIMINALÍSTICA

DOI: http://dx.doi.org/10.15260/rbc.v4i3.103

Prosodic characteristics of different varieties of Brazilian Portuguese

A.C. Constantini [a*], P.A. Barbosa [b]

[a] *Departamento de Desenvolvimento Humano e Reabilitação, Curso de Fonoaudiologia, FCM/UNICAMP, Campinas (SP), Brasil*
[b] *Instituto de Estudos da Linguagem, UNICAMP, Campinas (SP), Brasil*

* *Endereço de e-mail para correspondência: aconstantini@fcm.unicamp.br.*

Recebido em 17/09/2015; Revisado em 09/12/2015; Aceito em 10/12/2015

Resumo

Analisamos um conjunto particular de parâmetros prosódicos para caracterizar sete variedades do Português Brasileiro (PB). As amostras de fala de 35 indivíduos do sexo masculino, com idade média de 35,4 anos, de São Paulo, Minas Gerais, Rio de Janeiro, Paraná, Distrito Federal (DF), região Nordeste (NE) e região Norte (N) foram coletadas. As amostras foram segmentadas em unidades de vogal-a-vogal (unidades VV) e oito medidas foram extraídas: taxa de elocução, média, desvio padrão, assimetria de z-score suave de duração da unidade VV, taxa de proeminência, taxa de não-proeminente da unidade VV/s, ênfase espectral e mediana de F0. One-Way ANOVA ($p < 0,05$) indicou significado para a taxa de não-proeminente da unidade VV, o desvio padrão, a assimetria de z-score, a ênfase espectral e a mediana de F0. Os resultados separaram variedades faladas no Norte daquelas no Sul. A descrição das características prosódicas é relevante para fins forenses, já que pode ser usado como prova e apoio durante a resolução do caso.

Palavras-Chave: Fonética Forense; Dialetos; Prosódia.

Abstract

We analyzed a particular set of prosodic parameters in order to characterize seven Brazilian Portuguese (BP) varieties. Speech samples from 35 male subjects, mean age of 35.4 years, from São Paulo, Minas Gerais, Rio de Janeiro, Paraná, Distrito Federal (DF), Northeast region (NE) and North region (N) were collected. The samples were segmented into Vowel-to-vowel units (VV units) and eight measures were extracted: speech rate, mean, standard deviation and skewness of smooth z-score of VV unit duration, prominence rate, rate of non-prominent VV unit/s, spectral emphasis and median of F0. One-Way ANOVA ($p<0.05$) has indicated significance for rate of non-prominent VV unit/s, standard deviation and skewness of z-score, spectral emphasis and median of F0. The findings separated varieties spoken in the North from those spoken in the South. The description of prosodic characteristics is relevant for forensic purposes since it can be used as evidence and support during case resolution.

Keywords: Forensic Phonetics; Dialects, Prosody.

1. INTRODUCTION

This paper explores a particular set of prosodic characteristics of different varieties of Brazilian Portuguese (BP). Usually, prosody is not the first choice to study differences among dialects, where the main focus is on segmental features and lexical differences. Nevertheless, we show here that prosody can reveal differences and commonalities across dialects which can be of use in sociolinguistic research.

Prosody has been defined as the suprasegmental phonetic information present in spoken utterances and is usually studied from three classic parameters: duration, fundamental frequency and intensity [1]. The aspects of production and perception in speech prosody have direct effects on communication, since they contribute to the indication of a declarative statement or an interrogative statement, for example. Segmentation and organization of speech, expression of emotions and speaker recognition are other examples of speech prosody functions in communication.

The prosody function that is particularly significant for the field of forensic phonetics is speaker identification. One of the aspects that can identify a speaker is the prosody present in the dialect or variety, which he/she speaks. The description of prosodic characteristics in a speaker identification task can be the key for a case resolution when there is doubt about the

place of origin of the suspect, for example. In addition, the characterization of dialects becomes important in another type of task often performed in forensic linguistics: speaker profiling. Reference [2] defines speaker profiling as the task of isolating a specific group of subjects sharing certain properties such as dialect, speech style and foreign accent.

According to Reference [2], if the talker speaks a dialect with relatively few speakers and a telephone call that may be traced to an area where a different dialect is spoken, the number of possible perpetrators can sometimes be narrowed down to only a handful.

The acoustic phonetic analysis of speech is one of the methods used to carry out forensic speaker identification. A survey conducted by Reference [3] with thirty-six professionals on forensic speaker comparison revealed that the main suprasegmental phonetic information to be measured was the fundamental frequency and its statistical descriptors (mean, median, standard deviation, range, and others). Nevertheless, the experts emphasized that those aspects are used as an elimination tool rather than an identification tool. Reference [3] also pointed out that all experts analyze segmental features as formant frequencies and consonant-vowel transitions, among others.

There are several types of problems that can influence the extraction of segmental and suprasegmental measures in speech recordings during the speaker identification task, such as voice disguise, signal distortion and background noise. Prosody seems to assume an important role in speech recordings with background noise. Reference [4] studied the impact of background noise in speaker recognition. The authors had added white noise in telephone calls and tested the possible cues listeners use to identify speakers in the presence of noise. The experimental analysis has included the presentation of excerpts spanning 10 to 20 seconds to 90 listeners. The listeners' task was to correlate the noisy recordings with longer excerpts without noise (approximate duration of 90 seconds). A 7-level verbal scale was established ranging from "I am certain that the speakers are different" to "I am certain that the speakers are the same". At the end of the experiment, the listeners were questioned as to which parameters had motivated their decisions. The authors concluded that 30% of listeners considered the matching between the noisy telephone call and the speech recording without noise to be correlated to "pronunciation" and "accent". Ranking immediately after that, 20% of listeners used "timbre" to do the matches. Third in the rank, 18% of the answers were related to intonation. The authors concluded that the knowledge about accent and pronunciation of different varieties are notably useful on speaker recognition tasks.

1.1. Prosodic aspects on different varieties

As we showed before, prosody plays an important role in describing the characteristics of different varieties. However, most of the work in regional variation was done on segmental and lexical differences among varieties.

Reference [5] studied prosodic variability in different dialects of American English. The authors claimed that in the US, researchers are largely focused on segmental sources of variability. To try to fill this gap, the authors examined the effects of regional dialect and gender on prosodic patterns in two varieties of American English: Midland and Southern. The investigation was conducted based on the analysis of speech rate, pitch patterns and pause distribution on those two varieties and the results showed that speech rate is similar between the two varieties. The explanation for that could be related to the type of task the talkers were submitted to (reading task) and the geographic proximity of the varieties. In respect to pause distribution, the results showed the Southern male talkers had produced more pauses than the Southern female talkers or the male or female Midland talkers. The results also revealed the effects region and gender have on the frequency of pitch accent types and phrasal-boundary tones combinations. These effects were related to the preference for L*+H in females talkers from both dialects and a rejection of H* pitch accents relative to male talkers.

Some prosodic characteristics of two Brazilian varieties were studied by Reference [6]. The authors had focused on the rhythmic aspects of Minas Gerais (MG) variety and Bahia (BA) variety. Two women and two men from those varieties were recruited and asked to read a selected text 10 times in three different speech rates: fast, normal and slow. The results showed the two varieties had the same normal speech rate and that MG had a fast speech rate faster than that of BA.

The historical division of BP varieties that has been influencing the development of new work on linguistic variation is due to [7]. In this seminal work, it is suggested the use of a line crossing Brazilian territory and creating two major dialectal regions: North and South. North varieties, according to the description of Reference [7], include the states of Amazonas, Acre, Pará, part of Goiás and the Northeast territory up to Bahia, which is considered to be intermediate between North and South. South varieties contain part of Bahia territory, Minas Gerais (north, northwest and northeast of the state of Minas Gerais), part of Goiás territory, Rio de Janeiro, Espírito Santo and the South states: São Paulo, Paraná, Santa Catarina and Rio Grande do Sul.

Currently, it is clear that the geographical division proposed by [7] has many limitations related to the use of only two discriminant criteria: open/closed pronunciation

of pre-stressed vowels and "cadence" of speech. Nevertheless, his division is still thought of as a historical fact and is widely used by other authors as a source of comparison to new studies. Reference [8] investigated the linguistic atlas of Bahia [9] and Minas Gerais [10] and after that, the author confirmed division, at least for those specific regions [7].

Thus, the available division of Brazilian Portuguese varieties remains the same until today. However, it is expected that this situation will change after the publication of the Brazilian Linguistic Atlas (AliB – Atlas Linguístico Brasileiro). Establishing a new division is quite complex due to the size of the territory and the presence of descriptions of different varieties spoken within the same State (case of Minas Gerais) [11]. In the current study, we decided to adopt [7] proposal as the standard and, for research purposes, the selection of subjects per variety was done according to the State of origin, or region, as described in the next section, with exception of North and Northeast regions.

To adopt the political division of the States of the Federation as an approximation of the dialectal division in not contradictory to old dialectal divisions. In fact, statistical analyses can further help rethink the delimitation of the Brazilian territory in relation to prosodic aspects.

This paper aims to describe and characterize the prosodic parameters of seven varieties of Brazilian Portuguese.

2. METHODOLOGY

2.1. Corpora

Our corpus consists of records of informal interviews of 35 male subjects with mean age of 35.4 years, from seven different Brazilian state/regions. This corpus is a subset of a major corpus built by the Brazilian Federal Police, the Forensic Brazilian Corpus. It was built to contribute to research on forensic phonetics.

All 35 subjects are police officers from the following Brazilian state/regions (from now on simply referred as "regions"): São Paulo (SP), Rio de Janeiro (RJ), Minas Gerais (MG), Distrito Federal (DF), Paraná (PR), Northeast region (NE) and North Region (N). The distribution of subjects per region is described in Table 1. In order to be considered a representative talker of each region, we established as criteria that the subjects must be born and raised in that region.

The speech samples had duration of approximately five minutes and were recorded in a sounded-treated booth. The recordings consisted of interviews in which the subjects were encouraged to talk about familiar topics such as family, professional life and vacations. The subjects were allowed to talk as much as they wanted to and were hardly ever interrupted by the interviewer. The recordings can be considered spontaneous since all the subjects behaved in an informal way. The Brazilian Federal Police authorized the use of the recordings for the purposes of development of this research.

Table 1. Division of subjects per region.

Region	Number of subjects
São Paulo (SP)	6
Rio de Janeiro (RJ)	6
Minas Gerais (MG)	6
Região Nordeste (NE)	6
Distrito Federal (DF)	5
Paraná (PR)	4
Região Norte (N)	2
Total	35

With the exception of NE and N regions, all the subjects were selected according to their state of origin. The decision of grouping subjects from different States to compose the NE for example, was made based on Reference [7]' division of Brazilian dialects. This will allow verifying if the NE could still be considered as a large and homogeneous variety, at least in terms of prosody. This decision happened to be appropriate because this region presented a low standard deviation for almost all acoustic parameters. As regards the N region variety, it was difficult to find more subjects who could be included in the corpus according to our selection criteria. This is related to the low demographic density of the N region compared with the other Brazilian regions studied here. Further generalizations should be done carefully, since the results are related to the specific population studied in this research.

2.2. Measuring techniques

For the data analysis, each speech sample with approximately 5 minutes was divided into three small excerpts of approximately 100 seconds, which allowed for a better performance of the scripts used in the following stages. Only two samples were divided into four excerpts and one sample was divided into five. The total number of analyzed excerpts was 109.

At the next stage, all excerpts were segmented into Vowel-to-Vowel (VV) units. The VV units are phonetic syllables starting at the onset of a vowel and ending at the onset of the following vowel, including the consonants between them. The segmentation in VV units was performed automatically using the BeatExtractor Praat script [12, 13]. Early studies demonstrated that VV units

are minimal units of rhythmic processing and are also acoustic correlations of phrase stress [14, 15]. Therefore, the study of the syllabic duration, delimited according to VV units, could reveal particular characteristics of each subject and possibly allow grouping them into their varieties, which is the main hypothesis of the present research.

After the segmentation process, the excerpts were transcribed using an ASCII-based notation, which is exemplified in Figure 1. The segmentation in VV units can be seen in the bottom of Fig. 1. The first one is "ek(i)p" with an elided "i".

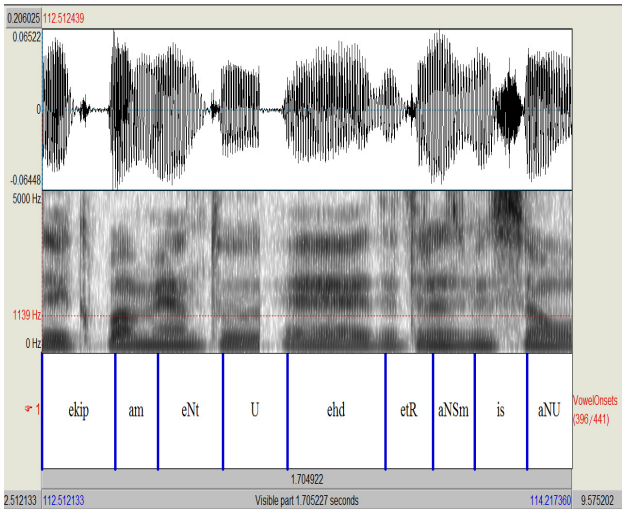

Figure 1. Excerpt of "equipamento é de transmissão" (transmission equipment). The top layer has the sound wave, the intermediate layer has the wideband spectrogram and the bottom layer has the standard notation ASC II transcription of the excerpt.

At the fourth stage, we used another script running onto Praat, named ProsodicDescriptorExtractor, which extracts 13 parameters of each excerpt that had been segmented into VV units and transcribed at the previous stages. For this, the ProsodicDescriptorExtractor [16] uses an audio file and an annotation file (Praat TextGrid). From the 13 extracted parameters, we have chosen eight, described in Table 2.

The eight parameters were selected considering acoustic measures that are frequently used for forensic phonetics purposes as spectral emphasis, speech rate and fundamental frequency. Moreover, [17] has used those same parameters with satisfactory results on individualization of prosodic characteristics and discrimination of differences in rhythmic structure across two different speech tasks: story-telling and reading.

In our study, the first parameter is speech rate (SR), measured in VV units per second. The majority of the other analyzed parameters consider the normalized z-scored duration. The z-score measure considers the normalized duration of VV units along the utterances using equation (1). The normalization technique allows signaling phrase stress positions throughout the utterance, which corresponds roughly to filtered (or smoothed) z-score local peaks.

$$z^{i}_{smooth.} = \frac{5.z^{i} + 3.z^{i-1} + 3.z^{i+1} + 1.z^{i-2} + 1.z^{i+2}}{13} \quad (1)$$

In (1) z^{i} is the z-score of the current VV duration, z^{i-1} is the z-score of the previous VV duration and z^{i+1} is the z-score of the following VV duration. The other indexes follow that reasoning accordingly.

The second, third and fourth parameters are respectively mean (mz), standard deviation (SDz) and skewness (Skz) of smoothed z-scored VV durations. According to Reference [15], these three measures reveal the structure of duration-related prominence degree and boundary strength, pooled together, throughout the excerpts.

The prominence rate (Pr) is measured by the rate of production of local peaks of smoothed z-score per second. The fundamental frequency (F0) is calculated by its median since from a statistical point of view this measure is considered robust and more resistant to outliers. Spectral emphasis (Emph) is the relative intensity concentrated at the high frequencies of the spectrum and it is measured by the difference between the intensity of the whole spectrum (0 to 11.000 Hz) and the intensity right above the fundamental frequency band (from 0 to 350 Hz) [18]. The last parameter is the rate of non-prominent VV units. This measure does not take the salient VV units into account when computing the rate of local peaks of smoothed z-scores.

2.3. Statistical Analyses

The software R [19] was used to conduct the statistical analysis. To investigate differences among the prosodic-acoustic measures in the studied varieties, the statistical test One-Way ANOVA was selected. The significance level adopted was 0.05 and the factor was named VARIETY. When a difference was revealed, two post-hoc tests were conducted in different times: Tukey Honestly Significant Difference (Tukey HSD) and Duncan, both with p-value <0.05. Tukey HSD test searches for statistically significant differences in paired groups after a significant One-Way ANOVA test. The Duncan test has a similar purpose but instead of testing for differences, it creates homogeneous groups from non-significant mean differences.

The statistical technique Linear Discriminant Analysis (LDA) was also applied to verify whether the prosodic-acoustic parameters could predict each variety. The LDA is a linear classifier based on the calculation, using Bayes

Theorem, of the posterior probabilities of each class (in the present case, the variety regions of Table 1) given a set of parameters (the prosodic features of Table 2). The set of parameters is classified in the class of maximum posterior probability. The likelihood functions (probabilities of parameters given class) are estimated assuming they are Gaussians of equal covariance matrix, with means and a priori probabilities simply taken as averages from the training subset. For more information about LDA see, for example, [20].

Table 2. Descriptions of the eight acoustic parameters analyzed using ProsodicDescriptorExtractor script. In the column Parameter are presented the abbreviations of each analyzed parameter as used in the script.

Parameter	Description
SR	Speech Rate (VV units/s)
Mz	Mean of smoothed z-score of VV unit duration
SDz	Standard deviation of smoothed z-score of VV unit duration
Skz	Skewness of smoothed z-score of VV unit duration
Pr	Prominence rate (peaks of z-score/s)
F0	Median of fundamental frequency
Emph	Spectral emphasis
Ur	Rate of non-prominent VV/s

3. RESULTS

Statistical analyses indicated significant difference for the following parameters: spectral emphasis, median of fundamental frequency, standard deviation of z-score, skewness of z-score and rate of non-prominent VV units/s (Table 3).

A note is necessary regarding standard deviation of z-score (SDz). Although it showed significance by One-Way ANOVA, the more conservative Tukey HSD test was not able to reveal statistically significant differences between varieties considering a p-value of 0.05. For this reason, for Tukey HSD analysis of SDz we took the p-value of 0.15 as an indication of a marginal statistical difference.

Table 3. Results for the prosodic-acoustic parameters. Values followed by (*) presented statistically significance.

Prosodic-acoustic Parameters	P-value
Spectral emphasis	<0,001*
Median of F0	<0,001*
Mean of z-score	0.2
Prominence rate	0.8
Standard deviation of z-score	0.01*
Skewness of z-score	0.04*
Speech rate	0.09
Rate of non-prominent VV units/s	0.02*

The results will be presented as follows in boxplots in an attempt to illustrate the findings obtained by One-Way ANOVA, Tukey HSD and Duncan test.

Spectral emphasis and median of F0 were able to distinguish a larger group of varieties. N Region and DF varieties exhibited the highest values of spectral emphasis (Figure 2) which separate them from the other varieties. However, there is no differentiation of spectral emphasis between DF and PR varieties. The reason for this finding is showed in Figure 2, which suggests that the mentioned similarity with DF is due to a high variability of spectral emphasis values presented in the speech of the PR variety.

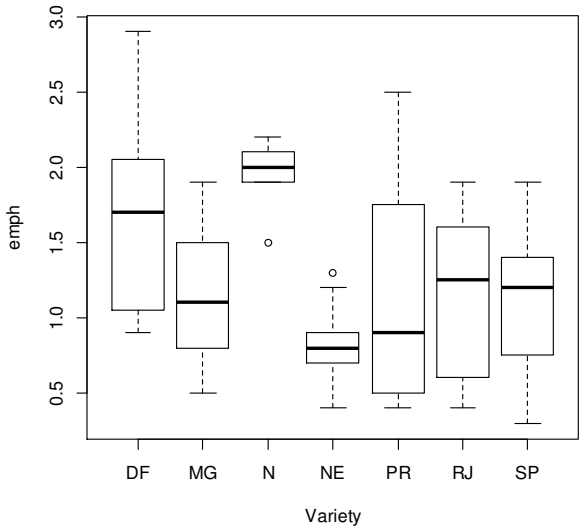

Figure 2. Boxplot showing spectral emphasis results in the studied varieties.

Median of fundamental frequency puts DF and N in a single group separating them from the other varieties (Figure 3). The obtained values for median of F0 for the subjects from N and DF are higher than those expected

for Brazilian men. Reference [21] reports a range of mean of F0 for men between 90 Hz and 150 Hz. Reference [22] found a mean of 127 Hz for male Brazilian talkers and 208 Hz for female Brazilian talkers. Considering the available data for fundamental frequency, the subjects analyzed in the current research are on the right edge of Reference [21] data and 20 Hz higher than the right edge of F0 in Reference [22]. Table 4 presents the values of mean and standard deviation of fundamental frequency median for all the studied varieties.

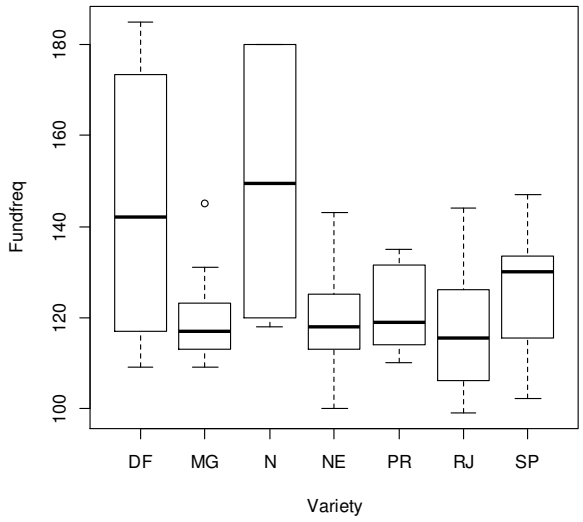

Figure 3. Boxplot showing median of fundamental frequency for the studied varieties.

Table 4. Mean and standard deviation of median of fundamental frequency for the studied varieties.

Variety	Median of F0	
	Mean (Hz)	Standard Deviation (Hz)
DF	144	29
MG	119	9
N	149	32
NE	119	11
PR	121	10
RJ	117	14
SP	125	13

While spectral emphasis and median of F0 were able to differentiate N and DF from all the other varieties, on the other hand, skewness of z-score and rate of non-prominent VV units/s have separated these two regions. Figure 4 illustrates the results for skewness of z-score.

The results for skewness of z-score set the N Region as the variety with the highest values of positive skewness and DF with lowest values of positive skewness. This result indicates that the N region speakers increase the duration of the VV units more often than DF speakers.

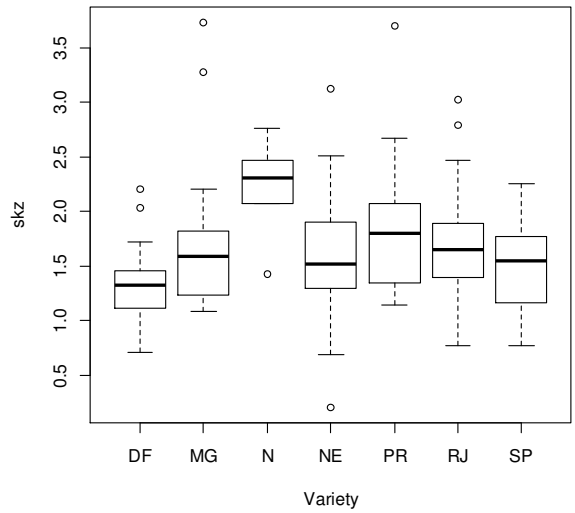

Figure 4. Boxplot showing skewness of z-score for the studied varieties.

The rate of non-prominent VV units also separated N and DF varieties. As Figure 5 shows, speakers from DF and SP exhibited a low rate of non-prominent VV units/s, while the N region speakers have a higher rate. It is important to highlight that the rate of non-prominent VV units is similar to articulation rate and this may indicate the speech from N has a higher pace than that of DF and SP.

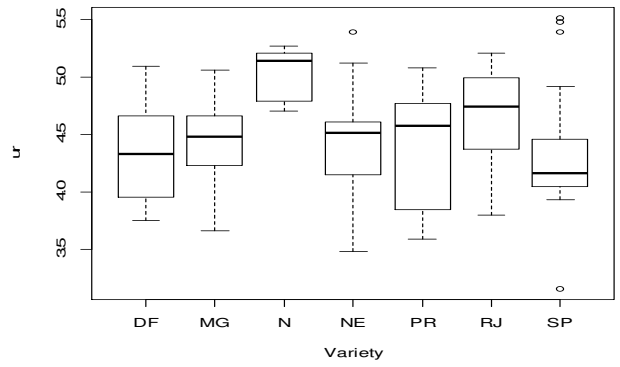

Figure 5. Boxplot showing rate of non-prominent VV units/s for the studied varieties.

As it has been said, standard deviation of z-score exhibited marginally significant results ($p<0.15$). Figure 6 illustrates the differences found in the standard deviation parameter. For the first time, two varieties from South (PR and SP) have been separated. Furthermore, the results indicate differences between the N and NE regions, as well as between DF and SP. The N region exhibits higher values of standard deviation than others.

The following results are related to the findings of the Duncan test. The results will be reported according to the groups created by the Duncan test. In some cases, it is possible to observe the formation of intermediate groups. The same parameters signal statistical significant differences in the two post-hoc tests, Tukey HSD and Duncan. We focus here on the results of the Duncan test to highlight grouping of varieties according to the parameters.

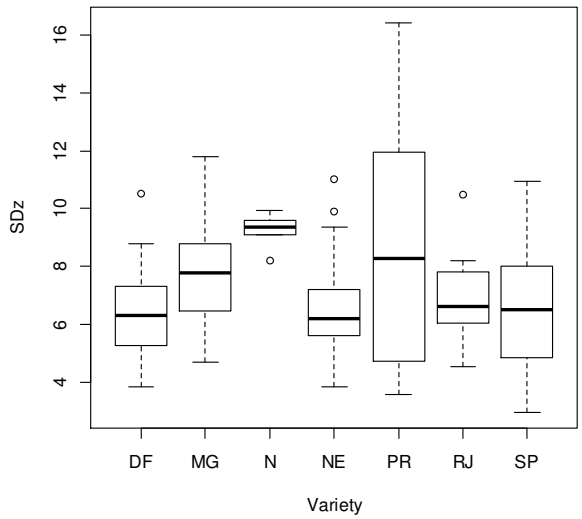

Figure 6. Boxplot showing standard deviation for the studied varieties.

The results of spectral emphasis are summarized in Table 5. The Duncan test results corroborate the Tukey HSD results: N and DF are still in the same group (group A) and the other varieties make up group B.

Table 5. Division of groups from Duncan test for spectral emphasis.

Spectral emphasis (dB)		
Groups	Varieties	Mean
A	N	1.95
A	DF	1.69
B	PR	1.16
B	MG	1.13
B	RJ	1.11
B	SP	1.1
B	NE	0.79

The same pattern was found for median of F0 (Table 6). Once more, the Duncan test corroborates the previous findings.

The standard deviation of z-score divides the varieties among two major groups (A and B) and one intermediate (AB) as shown in Table 7.

Table 6. Division of groups according Duncan test for median of fundamental frequency.

Median of F0 (Hz)		
Groups	Varieties	Mean
A	N	149.5
A	DF	144.6
B	PR	125.2
B	MG	121.8
B	RJ	119.9
B	SP	119.3
B	NE	117.3

Table 7. Division of groups from Duncan test for standard deviation of z-score.

Standard deviation of z-score		
Groups	Varieties	Mean
A	N	9.26
AB	PR	8.67
AB	MG	7.78
B	RJ	6.76
B	NE	6.59
B	DF	6.46
B	SP	6.44

Table 8. Division of groups from Duncan test for skewness of z-score.

Skewness of z-score		
Groups	Varieties	Mean
A	N	2.22
AB	PR	1.89
AB	MG	1.74
AB	RJ	1.73
AB	NE	1.61
B	SP	1.49
B	DF	1.36

The results presented in Table 7 indicate that the N region is the only variety representing group A, having the highest values of standard deviation. This finding may indicate N region talkers as the population having more variation in the duration of the VV units during their speech. PR and MG varieties represent the intermediate group AB. The designation of an intermediate group may suggest the results for those varieties are not too high to be included in group A and not too low to compose group B. On the other side, group B contains most of the varieties and is characterized by having low standard deviation of z-score.

Table 8 contains the results for skewness of z-score and once again, there is an intermediate group (AB) which is represented by four varieties. Group A has N region and group B has SP and DF. These findings corroborate the Tukey HSD test.

The last statistical significant parameter is the rate of non-prominent VV units/s, which is presented in Table 9. It shows N forming group A, RJ in the intermediate group AB and the other varieties in group B.

Table 9. Division of groups from Duncan test for rate of non-prominent VV units/s.

Rate of non-prominent VV units/s		
Groups	Varieties	Mean
A	N	2.22
AB	RJ	1.89
B	MG	1.74
B	NE	1.73
B	PR	1.61
B	SP	1.49
B	DF	1.36

Linear Discriminant Analysis (LDA) was used to predict the varieties from the results of the acoustic parameters. Initially, we took the set of eight parameters described in Table 2 to perform the LDA to predict each variety. Then, we used each parameter separately to perform a LDA aiming to predict the varieties from single parameters. The results for the set of parameters are presented in Table 10. No division between test and training sets was carried out. Our main goal is to test in a near future on a different material and to analyze the performance in the training set only, by now.

Table 10. LDA results of the correct proportion of predictions considering the set of parameters for the studied varieties.

Variety	Correct predictions
PR	75%
NE	66.60%
N	66.60%
DF	66.60%
RJ	50%
MG	38.20%
SP	31.60%

PR is the variety with the highest score of correct predictions, followed by NE region, N and DF, taking into account the set of eight selected parameters. SP variety had the lowest score by the LDA technique.

Table 11 presents the results obtained when the LDA was applied considering each parameter separately.

The use of LDA considering each parameter separately shows to be possible to discriminate correctly from two to four varieties, although the correct prediction has not been done with a high level of accuracy. The NE region variety was predicted by all the parameters with a minimum of 47% (for Mz parameter).

There is a difference in the results when the LDA technique was conducted considering the set of parameters in comparison to each parameter separately. The N region, for example, was not identifiable by any parameter separately but when the whole set of parameters was considered to make the predictions that variety exhibited 66.6% of correct predictions. This result for N region variety suggests the importance of considering the whole set of parameters to discriminate and characterize the varieties.

4. DISCUSSION

Knowledge about dialects has been playing an important role for forensic purposes [2, 23-24]. In forensic phonetics, the description of prosodic characteristics of varieties can be used as evidence supporting the guilty or innocent hypotheses. Therefore, the development of research that intends to describe and characterize different varieties is important in forensic linguistics.

In studies concerning different dialects or varieties from different languages, prosody has not been widely considered, since the main focus remains on phonological and lexical differences [5]. The same situation is valid for BP.

In the current study, we have proposed to measure eight prosodic-acoustic parameters to verify probable differences among the varieties of BP and five parameters have shown statistical significance: spectral emphasis, median of F0, skewness of z-score, rate of non-prominent VV units/s and standard-deviation of z-score. The results support the hypothesis that prosody, along with segmental features and lexical differences, is able to differentiate some of the varieties spoken in the Brazilian territory. From the five parameters, which have presented statistical significance, one is related to fundamental frequency (median of F0), another is related to intensity (spectral emphasis) and three are related to rhythmic structure (skewness and standard-deviation of z-score and rate of non-prominent VV units/s). This shows that the three classic parameters of prosody are involved in differentiating among BP varieties.

Table 11. LDA results of correct predictions considering the acoustic parameters separately.

Variety	Emph	F0	Mz	Pr	SDz	Skz	SR	ur
NE	85%	71%	47%	71%	85%	52%	66%	57%
SP	31%	57%	31%	0	5%	26%	31%	15%
N	0%	0%	0	0	0	0	0	0
DF	46%	53%	0	0	0	6%	0	0
MG	16%	0%	22%	0	44%	11%	0	0
PR	0%	0%	50%	0	33%	0	0	0
RJ	0%	16%	0	44%	0	0	11%	61%

Spectral emphasis and median of fundamental frequency reveal two main prosody-related groups: the N region and DF varieties are separated from all the others. On the other hand, skewness of z-score and rate of non-prominent VV units/s put the varieties in two different groups. We could claim that the similarities between N region and DF are related to the effects on F0 and spectral emphasis by a higher vocal effort during the speech and the differences are linked to rhythmic structure, since the N region has presented high values of positive skewness of z-score and high rate of non-prominent VV units/s. It is necessary to emphasize that DF variety is located at Midwest region, which is near to the N region. In addition, DF variety is considered a variety in formation, since, as a political division and a major urban area, the DF was founded in 1960 and has received a great amount of migrant from different Brazilian regions, resulting in an intense contact among different varieties. The results indicate that DF is similar to the N region in the use of intensity and fundamental frequency and similar to the South region in relation to rhythmic structure, which could be a result of the contacts between subjects of different varieties during DF foundation.

The decision to include different states from the NE region in the same group was appropriate. The findings indicate a low standard deviation (Figures 2, 3, 4, 5 and 6) for the NE variety and the LDA technique (Tables 10 and 11) suggests that individual acoustic parameters predict the NE variety with more accuracy than the other varieties.

In most cases, the NE variety exhibits lower values of standard deviation from all parameters, than those of the SP variety, which is composed by six talkers from the same state. The reason may be related to a likely prosodic distinction among the cities of origin of the subjects from SP state, which could be explored in future studies as an attempt to justify the high standard deviation found.

Varieties spoken more to the south as RJ, MG, PR and SP show few differences among each other. PR and SP varieties presented differences regarding standard deviation of z-score demonstrated by Tukey's test: SP showed less variation on the duration of VV units then PR. Duncan's test has created intermediate groups including some south varieties (Tables 7, 8 and 9) and this seems to reveal that South varieties exhibit differences from one another. However, those differences are not sufficiently strong to put the varieties into distinct groups.

It is important to consider what the corpora used on the current analysis reveals in the present. The findings for DF and N region could indicate that the' division proposed by Reference [7] is still valid, at least for the population studied in this work. On the other hand, for the NE region, which is also located in the North, it was not possible to find consistent differences between NE and the varieties in the South. One hypothesis is that the differences between NE and South varieties are more related to segmental and lexical distinctions.

5. CONCLUSION

The findings revealed prosodic distinctions involving the three classic prosodic parameters: duration, fundamental frequency and intensity. The differences showed here separate varieties spoken in the North from those spoken in the South. Considering the group of eight prosodic-acoustic parameters chosen to analyze the varieties, five of them were able to form distinct groups among the varieties. They are: spectral emphasis, median of F0, skewness of z-score, rate of non-prominent VV units/s and standard-deviation of z-score.

Spectral emphasis and median of F0 were the parameters responsible for the formation of the two major homogeneous groups: the DF variety and the N region variety, which are then separated from the others. On the other hand, N region and DF varieties were separated by parameters related to rhythmic structure.

The analysis of eight prosodic-acoustic parameters proved to be helpful in characterizing seven varieties of Brazilian Portuguese. The description of prosodic

characteristics of varieties is relevant for forensic purposes since it can be used as evidence and support during case resolution.

ACKNOWLEDGMENTS

We would like to acknowledge FAPESP, process 2011/00471-8, for the financial support during the development of this research and the Brazilian Federal Police for providing the speech sample recordings.

REFERÊNCES

[1] I. Lehiste. Suprasegmentals. Cambridge, Mass.: MIT Press.1970.

[2] A. Eriksson. The disguised voice: imitating accents or speech styles and impersonating individuals. In C. Llamas; D. Watt (Eds.). Language and Identities. Edinburgh: Edinburgh University Press, p. 86–96, 2010.

[3] E. Gold, P. French. International practices in forensic speaker comparison. *International Journal of Speech, Language & the Law* **18(2)**, 293-307, 2011.

[4] A. Alexander, F. Botti, D. Dessimoz, A. Drygajlo. The effect of mismatch recording conditions on human and automatic speaker recognition in forensic applications. *Forensic Sci. Int.* **146**, S95-S99, 2004.

[5] C.G. Clopper, R. Smiljanic. Effects of gender and regional dialect on prosodic patterns in American English. *J. Phon.* **39(2)**, 237–245, 2011.

[6] A.R Meireles, J.P. Tozetti, R.R. Borges. Speech rate and rhythmic in Brazilian Portuguese. In: Speech Prosody 2010 Conference, 2010, Chicago. Proceedings of the Speech Prosody 2010 Conference. Chicago: RG. **1**, 1-4, 2010.

[7] A. Nascentes. O linguajar carioca. Rio de Janeiro: Organização Simões, 1953.

[8] S.A.M. Cardoso. Tinha Nascentes razão? Considerações sobre a divisão dialetal do Brasil. *Estudos: Linguísticos e Literários* **5**, 47-59, 1986.

[9] J. Ribeiro. Esboço de um Atlas Linguístico de Minas Gerais. Rio de Janeiro: Ministério da Educação e Cultura - Casa de Rui Barbosa - Universidade Federal de Juiz de Fora, 1977.

[10] N. Rossi, C. Ferreira, D. Isensee. Atlas Prévio dos Falares Baianos. Rio de Janeiro: Ministério de Educação e Cultura - Instituto Nacional do Livro, 1963.

[11] E.F. Martins. Atlas linguístico do estado de minas gerais: O princípio da uniformidade da mudança linguística nas características fonéticas do português mineiro. *Revista Virtual de Estudos da Linguagem* **4(7)**, 1-13, 2006.

[12] P. Boersma, D. Weenink. Praat: Doing phonetics by computer. Version 5.3.39, 2013.

[13] P.A. Barbosa. Incursões em torno de ritmo da fala. Campinas: Editora Pontes, 2006

[14] B. Pompino-Marschall. The syllable as a prosodic unit and the so-called P-centre effect. Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München. **29**, 65-123, 1991.

[15] G. Dogil, G. Braun. The PIVOT model of speech parsing. Verlag der Österreichischen Akademie der Wissenschaften. Viena, 1988.

[16] P.A. Barbosa. ProsodicDescriptorExtractor. Script available with the author, 2011.

[17] P.A. Barbosa, W. Silva. A New Methodology for Comparing Speech Rhythm Structure between Utterances: Beyond Typological Approaches. In: Computational Processing of the Portuguese Language 10th International Conference, PROPOR 2012, Coimbra, Portugal, April 17-20, 2012.

[18] H. Traumüller, A. Eriksson. Acoustic effects of variation in vocal effort by men, women, and children. *J. Acoust. Soc. Am.* **107(6)**, 3438-3451, 2000.

[19] The R Foundation for Statistical Computing. Available at: www.r-project.com. Version 3.2.0.

[20] M.J. Crawley. The R Book. 2nd edition. Ed. Wiley. **5**, 744, 2013.

[21] M. Behlau. Voz: o livro do especialista. Rio de Janeiro. Editora Revinter. p. 86-176, 2001.

[22] Z. Camargo. Voice quality and gender: some insights on correlation between perceptual and acoustic dimensions In: 6th International Conference on Speech Prosody, 2012, Shanghai. Abstract Book Speech Prosody 2012. Shangai: Tongji University Press, p.115-118, 2012.

[23] W.R. Shuy. Dialect as evidence in law cases. *J. English Linguistics* **23**, 1990-1995, 1990.

[24] S. Moosmüller. Phonological variation in speaker identification. *International Journal of Speech Language and the Law* **4(1)**, 29-47, 1997.