



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE ESTUDOS DA LINGUAGEM
Bacharelado em Linguística



LUÍS HENRIQUE GOMES ALVARENGA

Variação prosódica dialetal do português brasileiro: nova forma de recrutamento

Trabalho de Conclusão de Curso de graduação
apresentado como requisito parcial para obtenção do
título de Bacharel em Linguística pela Universidade
Estadual de Campinas.

Orientador: Prof. Dr. Plínio Almeida Barbosa

LUÍS HENRIQUE GOMES ALVARENGA

Variação prosódica dialetal do português brasileiro: nova forma de recrutamento

Trabalho de Conclusão de Curso de graduação
apresentado como requisito parcial para obtenção do
título de Bacharel em Linguística pela Universidade
Estadual de Campinas.

Orientador: Prof. Dr. Plínio Almeida Barbosa

Campinas
2023

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Estudos da Linguagem
Ana Lúcia Siqueira Silva - CRB 8/7956

Al86v Alvarenga, Luís Henrique Gomes, 2001-
Variação prosódica dialetal do português brasileiro : nova forma de
recrutamento / Luís Henrique Gomes Alvarenga. – Campinas, SP : [s.n.], 2023.

Orientador: Plínio Almeida Barbosa.
Trabalho de Conclusão de Curso (graduação) – Universidade Estadual de
Campinas, Instituto de Estudos da Linguagem.

1. Prosódia. 2. Variação. I. Barbosa, Plínio Almeida, 1966-. II. Universidade
Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.

Informações adicionais, complementares

Título em outro idioma: Dialectal prosodic variation in Brazilian Portuguese: new recruitment method

Palavras-chave em inglês:

Prosody
Variation

Titulação: Bacharel em Linguística

Banca examinadora:

Plínio Almeida Barbosa (orientador)
Renata Regina Passetti
Ana Carolina Constantini

Data de entrega do trabalho definitivo: 07-12-2023

*“Só se vê bem com o coração.
O essencial é invisível aos olhos.”*

Antoine de Saint-Exupéry

AGRADECIMENTOS

A princípio, agradeço ao meu orientador, o Prof. Plínio Barbosa, uma pessoa de sabedoria e bondade admiráveis. Sou muito grato por todos seus ensinamentos e oportunidades que tive ao participar do grupo de pesquisa de prosódia da fala.

Agradeço à minha família que sempre me apoiou e torceu por mim em todas as minhas conquistas, mesmo estando distantes.

A todos os meus amigos que estiveram ao meu lado durante o período de escrita deste trabalho e me ouviram falar incessantemente sobre ele, a nossa convivência me ajudou a amadurecer e manter a calma mesmo em momentos de ansiedade.

Aos Insurgentes, Ana, Juliana, Amanda e Guilherme, que foram muito importantes para tornar o período da pandemia mais tolerável e agradável em meio a tantas dificuldades.

Aos professores do curso de linguística que sempre trouxeram muita erudição em suas disciplinas e ajudaram muito na minha formação como linguista.

Aos participantes desta pesquisa, pois sem sua contribuição, não seria possível a produção deste trabalho.

Por fim, agradeço à Unicamp por todas as oportunidades, experiências de vida e amizades que fiz durante esse período da graduação. Estudar nesta universidade me fez ser uma pessoa melhor.

RESUMO

O presente trabalho tem como objetivo analisar e discutir sobre a variação prosódica dialetal no português brasileiro, visto que há uma carência de trabalhos sobre dialetos em seus aspectos prosódicos no Brasil. Assim, foram coletadas 43 gravações de 22 sujeitos de capitais de sete estados brasileiros: um homem da Bahia, um homem do Espírito Santo, dois homens e duas mulheres de Minas Gerais, uma mulher e um homem do Piauí, três mulheres e dois homens do Rio de Janeiro, uma mulher de Santa Catarina e três homens e cinco mulheres de São Paulo. Devido a questões estatísticas, analisou-se apenas áudios de 17 falantes de três capitais brasileiras (Belo Horizonte, Rio de Janeiro e São Paulo) com idade entre 20 e 30 anos. Desses áudios foram retirados e analisados quatorze parâmetros prosódico-acústicos, dentre os quais dez são relacionados à entonação (mediana da frequência fundamental ou f_0 , semi-amplitude entre os quartis de f_0 , valores mínimos de f_0 , valores máximos de f_0 , desvio padrão dos picos de f_0 , desvio padrão dos picos de f_0 no eixo temporal, primeira derivada positiva da média de f_0 , primeira derivada negativa da média de f_0 , desvio padrão da primeira derivada positiva de f_0 , desvio padrão da primeira derivada negativa de f_0), dois à intensidade (Ênfase espectral e o coeficiente de variação da intensidade) e dois ao ritmo (Taxa de elocução e taxa de articulação). A partir disso, este trabalho se propôs a analisar quais dos parâmetros prosódicos supracitados são capazes de diferenciar e caracterizar as variedades analisadas. Para a análise estatística foram feitos testes de variância ANOVA para verificar as diferenças entre os estados, assim como um modelo de regressão de efeitos mistos com os sujeitos sendo a variável aleatória, para verificar o efeito dos sujeitos sobre a amostra. Os resultados mostram que as três variedades se diferenciam em relação a diferentes parâmetros, sendo a mediana de f_0 o parâmetro que diferenciou as mulheres de São Paulo das outras variedades. Já a semi-amplitude entre os quartis de f_0 diferenciou os falantes de Minas Gerais dos outros e, de forma semelhante, as taxas de elocução e articulação também diferenciam Minas Gerais das outras variedades.

Palavras-chave: Prosódia; variação dialetal; descritores acústicos; taxa de elocução.

ABSTRACT

This study aims to analyze and discuss dialectal prosodic variation in Brazilian Portuguese, given the lack of research on dialects in their prosodic aspects in Brazil. Thus, 43 recordings of 22 subjects from the capitals of seven Brazilian states were collected: one man from Bahia, one man from Espírito Santo, two men and two women from Minas Gerais, one woman and one man from Piauí, three women and two men from Rio de Janeiro, one woman from Santa Catarina, and three men and five women from São Paulo. Due to statistical considerations, audio recordings of only 17 speakers from three Brazilian capitals (Belo Horizonte, Rio de Janeiro, and São Paulo) aged between 20 and 30 were analyzed. From these recordings, fourteen prosodic-acoustic parameters were extracted and analyzed, including ten related to intonation (median fundamental frequency or f_0 , semi-amplitude between f_0 quartiles, minimum f_0 values, maximum f_0 values, standard deviation of f_0 peaks, standard deviation of f_0 peaks on the temporal axis, positive first derivative of f_0 mean, negative first derivative of f_0 mean, standard deviation of positive first derivative of f_0 , standard deviation of negative first derivative of f_0), two related to intensity (spectral emphasis and coefficient of intensity variation), and two related to rhythm (speech rate and articulation rate). From this, the study aimed to analyze which of the aforementioned prosodic parameters are capable of differentiating and characterizing the analyzed varieties. For statistical analysis, ANOVA tests were conducted to verify differences between states, as well as a mixed-effects regression model with subjects as the random variable to assess the subjects' effect on the sample. The results indicate that the three varieties differ in relation to different parameters, with the median f_0 being the parameter that differentiated São Paulo women from the other varieties. Conversely, the semi-amplitude between f_0 quartiles differentiated speakers from Minas Gerais from the others, and similarly, speech and articulation rates also differentiated Minas Gerais from the other varieties.

Key-words: Prosody; dialectal variation; acoustic descriptors; speech rate.

LISTA DE FIGURAS

Figura 1: Proposta de divisão dialetal do território brasileiro feita por Nascentes (1950)	13
Figura 2: Segmentação em trechos do texto "O monge desastrado": A parte mais acima é a representação do áudio em ondas sonoras. Abaixo, tem-se, respectivamente, o espectrograma, a transcrição ortográfica, o trecho de fala completo, a marcação de pausas silenciosas e, por fim, a quantidade de sílabas fonológicas presentes no trecho.	18
Figura 3: Boxplot da mediana de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP	22
Figura 4: Boxplot da Semi-amplitude entre os quartis de f_0 (SAQ) em Hertz de mulheres e homens dos estados de MG, RJ e SP	22
Figura 5: Boxplot dos valores mínimos de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP	23
Figura 6: Boxplot dos valores máximos de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP	23
Figura 7: Boxplot do desvio padrão dos picos de f_0 em picos por segundo de mulheres e homens dos estados de MG, RJ e SP	25
Figura 8: Boxplot do desvio padrão dos picos de f_0 no eixo temporal em picos por segundo de mulheres e homens dos estados de MG, RJ e SP	25
Figura 9: Boxplot das derivadas positivas de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP	26
Figura 10: Boxplot do desvio padrão das derivadas positivas de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP	26
Figura 11: Boxplot das derivadas negativas de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP	27
Figura 12: Boxplot do desvio padrão das derivadas negativas de f_0 em Hertz De mulheres e homens dos estados de MG, RJ e SP	27
Figura 13: Boxplot da ênfase espectral em decibéis para mulheres e homens dos estados de MG, RJ e SP	28
Figura 14: Boxplot do coeficiente de variação da intensidade em decibéis de mulheres e homens dos estados de MG, RJ e SP	28
Figura 15: Boxplot da taxa de elocução em sílabas por segundo de mulheres e homens dos estados de MG, RJ e SP	30
Figura 16: Boxplot da taxa de articulação em sílabas por segundo de mulheres e homens dos estados de MG, RJ e SP	30

LISTA DE TABELAS

Tabela 1: Descrição dos parâmetros prosódicos obtidos pelo <i>Prosody Descriptor Extractor</i>	19
Tabela 2: Valores da mediana e Semi-amplitude entre os quartis de f_0 (SAQ) em Hertz de mulheres e homens dos estados de MG, RJ e SP	21
Tabela 3: Mediana dos valores mínimos e máximos de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP	23
Tabela 4: Mediana dos valores de desvio padrão dos picos de f_0 em picos por segundo de mulheres e homens dos estados de MG, RJ e SP	24
Tabela 5: Mediana e desvio padrão das derivadas positivas de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP	26
Tabela 6: Mediana e desvio padrão das derivadas negativas de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP	27
Tabela 7: Valores da ênfase espectral (dB) e coeficiente de variação da intensidade em porcentagem de mulheres e homens dos estados de MG, RJ e SP	28
Tabela 8: Medianas das taxas de elocução e articulação em sílabas por segundo de mulheres e homens dos estados de MG, RJ e SP	30

SUMÁRIO

1. Introdução	9
2. Revisão da literatura	10
2.1. Prosódia	10
2.2. Estudos prosódicos dialetais no exterior	12
2.3. Estudos prosódicos dialetais no Brasil	13
2.4. Estilo	16
3. Objetivos e hipóteses	16
4. Metodologia	17
4.1. Corpus	17
4.2. Novo método de gravação	17
4.3. Transcrição dos dados	18
4.4. Parâmetros prosódico-acústicos analisados	19
4.5. Considerações da análise	20
4.6. Análise estatística	20
5. Discussão dos resultados	21
5.1. Entonação	21
5.2. Intensidade	27
5.3. Ritmo	29
6. Conclusão	30
7. Bibliografia	32
ANEXOS	34

1. Introdução

Estudos sobre variação dialetal têm como objetivo identificar e observar quais aspectos linguísticos são característicos das diferentes variedades presentes em uma língua. Nesse sentido, pode-se analisar características lexicais, sintáticas, e fonéticas, dentre outras, que estejam em variação na língua. Um exemplo desse tipo de trabalho supracitado reside na obra de Labov (1966) sobre a estratificação do /R/ em lojas de departamento no inglês nova-iorquino. O autor avaliou a distribuição da produção ou não do /R/ em posição de coda silábica em três lojas de departamento, com o intuito de avaliar a influência de classe social na língua. Mesmo que esse trabalho não tenha o intuito de comparar diferentes variedades do inglês estadunidense, Labov trouxe informações de características dos falantes daquela região de Nova York.

As pesquisas nessa área são bastantes frutíferas, pois em um país como o Brasil, que tem uma grande extensão geográfica, espera-se que haja variação em algum nível linguístico, principalmente fonético, devido à diversidade de sotaques encontrados no território nacional. Valendo-se dessa diversidade linguística, pode-se citar o projeto do Atlas linguístico do Brasil (AliB) que traz em sua análise um caráter mais comparativo entre dialetos. O projeto conta com a colaboração de diversas universidades brasileiras com o objetivo de criar um atlas geral do português brasileiro. Para a obtenção dos dados é aplicado um questionário e depois os participantes lêem uma frase com as características que se quer analisar. O questionário contém questões fonético-fonológicas, semânticas, lexicais e morfossintáticas. Em relação à fonética e fonologia, são aplicadas 160 questões, das quais 11 são sobre prosódia, mais especificamente entonação. Eles avaliam as diferenças entre sentenças interrogativas e assertivas, como também três expressividades da entonação: desagrado, contentamento e ordem (Cunha, 2006).

As instruções dadas aos participantes seguem a seguinte estrutura:

*Instrução: “Você / o(a) senhor(a) quer dizer a algumas pessoas que estão presentes que você / o(a) senhor(a) **está muito feliz com o resultado do trabalho**. Como é que você / o(a) senhor(a) diz?”.*

Resposta esperada: “Oh, gente, estou muito feliz com o resultado do trabalho.”

Embora a prosódia tenha sido tratada de uma maneira bem restrita, devido à natureza dos dados com um nível de controle muito alto, o que não é o ideal para uma análise acústica

da fala, ela já traz indícios das diferenças dialetais existentes. Atkinson (1968), em seu estudo, aponta para a importância da prosódia na diferenciação e identificação das línguas. A autora analisa a capacidade dos falantes de inglês em diferenciar o espanhol e o inglês apenas escutando a informação prosódica. Após os testes estatísticos, a autora consegue provar que a foi possível reconhecer as duas línguas sem a informação lexical e sintática. Partindo disso, é possível expandir esse tipo de análise para diferentes variedades de uma mesma língua, tanto através de um viés da percepção, como também o da produção.

Tendo isso em vista, o presente trabalho pretende preencher uma pequena parte da lacuna dos estudos dialetais e assim avaliar como a prosódia se comporta em diferentes variedades do português brasileiro. Haja vista de que existe uma carência de pesquisas na área; pois, em sua maioria, as pesquisas sobre variação linguística focam principalmente em características segmentais da fonética e lexicais, relegando a prosódia da língua a um segundo plano. Para o fim da análise, foram escolhidos treze parâmetros prosódico-acústicos de gravações de 17 falantes de três capitais da região Sudeste do Brasil (São Paulo, Rio de Janeiro e Belo Horizonte). Dos treze parâmetros acústicos, que serão apresentados na seção 4.3, dois são sobre ritmo, dois são sobre intensidade, enquanto os nove restantes são relacionados à entoação.

Este trabalho se divide em seis seções. A seção 2 é dividida em quatro sub-seções trazendo uma revisão sobre a literatura existente na área. Assim, respectivamente, traz-se a definição de prosódia e a descrição de seus correlatos acústicos, em seguida há uma revisão dos trabalhos sobre variação prosódica dialetal em diferentes línguas, por fim busca-se trazer uma definição de estilo que abarque os dados deste trabalho. Na seção 3 encontram-se os objetivos e hipóteses que norteiam as análises feitas. A seção 4 apresenta a descrição dos materiais e metodologias usadas neste estudo. Nele há as informações dos participantes, das gravações analisadas, como foram coletados os parâmetros acústicos e a análise estatística. Os resultados da análise estatística são apresentados na seção 5 junto às suas discussões. Por fim, este trabalho encerra-se na seção 6 com as conclusões principais e levantamentos para pesquisas futuras.

2. Revisão da literatura

2.1. Prosódia

A prosódia faz parte do interesse das áreas da Fonética e Fonologia; entretanto seu estudo não tem interesse em análises segmentais, mas sim na entoação e no ritmo da fala de

uma língua. O termo “Prosódia” vem do grego antigo “*prosōdias*” que significa em “sintonia com o canto”, esse termo foi usado pela primeira vez na República de Platão para fazer oposição entre o que foi dito e a forma como foi dita (Barbosa, 2019). Segundo Fletcher (2010), existem duas concepções de prosódia, uma relacionada à área da fonética e outra à da fonologia:

A prosódia ou características prosódicas são para muitos foneticistas e cientistas da fala sinônimos de variações em parâmetros suprasegmentais, como duração, intensidade e f_0 , que contribuem em várias combinações para a produção e percepção de ênfase, ritmo, cadência, acento lexical e entonação de um enunciado. [...] Contudo é impossível excluir efeitos de duração sub-silábica de qualquer discussão sobre ritmo acentual. [...] Isso nos leva ao uso fonológico mais abstrato do termo prosódia: como a(s) estrutura(s) fonológica(s) hierárquica(s) abstrata(s) de um enunciado e as relações de proeminência dentro dessa estrutura.¹ (Fletcher, 2010, p. 521).

Diante dessas definições, a da fonética, tida como a análise de características físicas e acústicas da fala, é a mais adequada para os objetivos deste estudo. Assim, para realizar uma análise prosódica, pode-se avaliar três correlatos físicos da fala de forma conjunta ou separada: A frequência fundamental (doravante f_0), a duração e a intensidade. A frequência fundamental está relacionada à vibração das pregas vocais, ou seja, corresponde a quantas vezes as pregas vocais oscilam em um segundo, sendo medida em Hertz (Hz) ou em semitons. Dessa forma, quanto maior o valor de f_0 mais aguda tende ser a voz e quanto menor o valor mais grave tende ser a voz do falante. A frequência fundamental pode ser controlada pelo falante para dar entoação ao enunciado e veicular um sentido ao que foi dito. Por exemplo, as sentenças interrogativas totais em português brasileiro, em que há uma subida melódica na palavra ou sílaba tônica sendo seguida de uma descida e por vezes maior saliência acústica da última palavra do enunciado (Moraes, 1998).

A duração, como o nome já diz, é a medida temporal de algum segmento, podendo ser linguístico, como as sílabas e trechos de fala, ou não, como as pausas e o ciclo respiratório. Em relação à sílaba, suas principais unidades são as taxas de elocução e de articulação, que podem ser expressas em palavras por minuto ou sílabas por segundo, ou seja, quantos desses elementos linguísticos são produzidos em um período de tempo (Fletcher, 2010). A diferença das duas taxas reside na inclusão ou não das pausas produzidas pelo falante, assim a taxa de elocução inclui as pausas em seu cálculo, enquanto a taxa de articulação não.

¹ Prosody or prosodic features are for many phoneticians and speech scientists synonymous with variations in suprasegmental parameters such as duration, intensity, and f_0 that contribute in various combinations to the production and perception of stress, rhythm and tempo, lexical tone, and intonation of an utterance. [...] However it is quite impossible to exclude sub syllabic timing effects from any discussion of stress [...]. This leads us to the more abstract phonological use of the term prosody: as the abstract hierarchical phonological structure(s) of an utterance, and prominence relations within that structure. (Tradução própria)

É importante ressaltar que a taxa de elocução geralmente é explicada de forma simplificada como "velocidade de fala", o que pode levar a uma interpretação equivocada, sugerindo que os articuladores da fala estão se movendo mais rapidamente. No entanto, o que ocorre é a fusão de alguns segmentos e a não produção de outros, permitindo a produção de mais unidades em um período de tempo mais curto. Portanto, é preferível utilizar o termo "taxa de elocução" em vez de "velocidade de fala" para descrever esse fenômeno (Barbosa, 2019).

Por fim, a intensidade é uma medida da força do som, ela é expressa em decibéis (dB), uma unidade que facilita a manipulação numérica e aproxima-se da sensação de volume “alto” e “baixo” experimentada pelos ouvintes. Seus descritores acústicos são referentes ao esforço vocal, ou seja, o nível de tensão usado nas pregas vocais.

Com isso em vista, percebe-se que as pesquisas sobre prosódia tornam-se bastante abrangentes, pois com apenas esses três correlatos físicos da fala é possível analisar desde a qualidade de voz de um falante, até diferenças de entoação entre sujeitos. Dentro dessa gama de possibilidades, pode-se avaliar quais dessas características prosódicas podem se portar como um fator para a diferenciação dialetal, haja vista, a priori, todas as concepções impressionísticas de que um dialeto é “arrastado” ou “cantado”. Concepções estas que serão melhor explicadas na seção 2.3.

2.2. Estudos prosódicos dialetais no exterior

Biadsky e Hirschberg (2009), com o auxílio de um sistema de reconhecimento de fala automático, conseguiram diferenciar quatro variedades do árabe baseando-se na prosódia: o árabe iraquiano, o árabe levantino, o árabe egípcio e o árabe do Golfo. Seu *corpus* é constituído por dados de fala de telefone dessas quatro variedades e foram analisados os dois primeiros minutos de cada gravação. Após realizar testes estatísticos, os autores encontraram diferenças relativas à entoação, à taxa de elocução e também à duração de vogais. Quanto ao reconhecimento automático dessas variedades, eles obtiveram uma precisão de 86.33% de identificação.

Ainda no tópico de reconhecimento automático da fala, pode-se citar o trabalho de Etman & Louis (2015). Nesse estudo foram analisadas características fonotáticas e prosódicas com o objetivo de identificar e diferenciar variedades do inglês estadunidense. Os autores realizaram testes de identificação com essas características de forma isolada e notaram, no entanto, que ao combiná-las a precisão de identificação do sistema aumentou para 83.3%, o que corroborou com uma de suas hipóteses de que adicionar dados prosódicos ao sistema

aumentaria sua precisão, mesmo com poucos dados de treinamento escassos. Em seguida, ao observar os resultados do sistema, puderam diferenciar o dialeto do Sul de outros quatro: o *South-Midland*, o de Nova York, o Ocidental e o do Norte.

2.3. Estudos prosódicos dialetais no Brasil

Em estudos pioneiros da área, por mais que a prosódia ainda não fosse o foco, a melodia da língua já era discutida de uma forma incipiente. Assim, pode-se destacar o trabalho de Antenor Nascentes (1953), em que o filólogo elaborou um dicionário do dialeto carioca da época, mas também discorreu brevemente sobre as diferenças linguísticas impressionísticas encontradas no Brasil. O autor foi o primeiro a propor uma divisão geográfica dos dialetos brasileiros para abarcar as diferenças melódicas encontradas na época. A princípio, ele separou o Brasil em duas macrorregiões: “os falares do Norte e os falares do Sul”, descrevendo a fala dessas áreas como “a fala *cantada* do nortista e a fala *descansada* do sulista”, utilizando como critérios a cadência e a abertura das vogais /e/ e /o/ no contexto pretônico. O primeiro adjetivo poderia se referir à melodia da fala, possivelmente sua variação na fala, enquanto que o segundo poderia se referir a uma taxa de elocução mais lenta.

Logo após sua macrodivisão, Nascentes separou o território do país em seis subfalares: O Amazônico, o Nordestino (pertencentes ao falar do Norte), o Baiano, o Fluminense, o Mineiro e o Sulista, (pertencentes aos falares do sul). Para elaborar essa divisão, o autor viajou por esses territórios e anotou suas percepções acerca das diferentes variedades encontradas. A representação cartográfica encontra-se na figura 1.

Figura 1: Proposta de divisão dialetal do território brasileiro feita por Nascentes (1950).



Outro trabalho com objetivo semelhante, mas anterior ao de Nascentes é o de Amadeu Amaral (1920) que visou descrever o dialeto da cidade de São Paulo na época denominado como caipira, analisando diferentes níveis linguísticos, como fonética, sintaxe e léxico para a criação de um dicionário. Em sua primeira descrição fonética, o autor apresenta ao leitor, de forma breve, aspectos da prosódia caipira:

Antes de tudo, deve notar-se que a prosódia caipira (tomando o termo prosódia numa acepção lata, que também abranja o ritmo e musicalidade da linguagem) difere essencialmente da portuguesa. O tom geral do frasear é lento, plano e igual, sem a variedade de inflexões, de andamentos e esfumaturas que enriquece a expressão das emoções na pronúncia portuguesa. (Amaral, 1920, p. 9).

Ao observar os trabalhos supracitados torna-se evidente que, mesmo de forma ainda embrionária, a prosódia já desempenhava um papel significativo na percepção das variedades estudadas. Essas análises ressaltam a importância de investigar mais a fundo a prosódia para uma compreensão mais abrangente da variação linguística.

Em estudos mais recentes, os pesquisadores já colocam a prosódia em primeiro plano, tratando e analisando seus correlatos acústicos como fatores para diferenciação dialetal, como é o caso do trabalho feito por Constantini (2014) que analisou uma amostra de fala enviada pela Polícia Federal do Distrito Federal de 35 sujeitos de cinco estados e duas regiões Brasileiras: Distrito Federal, Minas Gerais, Região norte (N), Região nordeste (NE), Paraná, Rio de Janeiro e São Paulo. Vale ressaltar que essa divisão feita em regiões se deve ao fato de que havia poucos sujeitos por estado e de que os sujeitos dessas regiões comportaram-se de forma semelhante, explicando assim o agrupamento.

Ao analisar os dados, a autora foi capaz de separar essas regiões em diferentes grupos, que na maioria das vezes foram o Distrito Federal e a Região Norte em um grupo separado do restante das regiões. As diferenças principais encontram-se nos parâmetros da mediana de f_0 , de esforço vocal e taxas de unidades silábicas de vogal a vogal (VV; Barbosa, 2006 *apud* Constantini, 2014). A autora também pôde avaliar que a região Norte e o Distrito Federal foram os grupos que tiveram os maiores valores de desvio-padrão de f_0 . A partir desses resultados, Constantini declara que a divisão feita por Nascentes (1953) ainda é aplicável, pelo menos em seu *corpus*, pois os correlatos prosódicos que demonstraram resultados estatisticamente diferentes distinguiram as variedades do Norte das variedades do Sul.

É possível também avaliar essas diferenças por um viés perceptivo, como é o caso do trabalho de Peres (2011), em que o autor fez a análise da percepção prosódica de três diferentes variedades brasileiras: o da Mooca em São Paulo, o de Senador Pompeu (CE) e o

de Pelotas (RS). O autor conduziu testes de percepção com 39 participantes, sendo 13 ouvintes de cada uma das localidades mencionadas. No primeiro teste, os estímulos apresentados aos participantes não continham elementos lexicais, permitindo apenas que a informação prosódica fosse ouvida. No segundo teste, os áudios foram manipulados para eliminar a variação melódica e torná-los monótonos.

De maneira geral, os resultados mostraram que no segundo teste houve um índice de acerto de 75%, enquanto no primeiro teste, o índice de acerto foi de 62%. Esses resultados sugerem que, mesmo na ausência de variação melódica significativa, os ouvintes ainda se baseiam em informações lexicais e fonéticas para identificar os dialetos. Além disso, o estudo de Peres indicou que os falantes identificaram com maior frequência o seu próprio dialeto em comparação com os dialetos das outras regiões em ambos os experimentos. O autor também fez uma análise dos parâmetros acústicos da produção em seu *corpus*, a variedade de São Paulo foi a que mais se diferenciou das outras. Já os parâmetros que se demonstraram significativos na análise estatística foram a mediana e o desvio padrão de f_0 .

Na sociolinguística a percepção tem um viés mais social, como a análise de estereótipos e meta-comentários dos falantes acerca da língua (Labov, 2008 [1972]). Baseando-se nisso, Freitag *et al* (2016) fizeram diferentes testes de percepção com universitários acerca do português brasileiro. Em sua breve apresentação sobre questões prosódicas, as autoras relatam que uma das características mais salientes apontadas pelos falantes é a taxa de elocução, trazendo adjetivos como “arrastado”, “cantado” e “lento” como característica das variedades do Nordeste; e “rápido” para variedades do Sul. Juntamente a essas características, as autoras também mostram dados de outras pesquisas para apoiar suas asserções, como a de Meireles e Gambarini (2011 *apud* Freitag *et al*, 2016) sobre as taxas de elocução de diferentes estados: Minas Gerais (6,3 S/s); São Paulo (6,6 S/s); Bahia (4,2 S/s) e Espírito Santo (4,0 S/s).

Por fim, outro projeto relevante a ser mencionado é o AMPER-POR, que tem como objetivo descrever a variação nas curvas melódicas de diferentes países falantes de línguas românicas, incluindo o português. De forma semelhante ao projeto ALiB, esse projeto analisa *corpora* compostos por áudios de leituras de frases interrogativas e assertivas; e as características fonéticas analisadas são mapeadas em seu site de forma interativa². No entanto, embora o estilo de leitura de frases ofereça a vantagem de uma maior comparabilidade entre os participantes, ele não reflete um estilo de fala natural, o que não é o mais adequado para pesquisas que buscam obter dados minimamente próximos da fala espontânea (Cunha, 2006).

² Disponível em: <http://www.varialing.eu/?page_id=272>

2.4. Estilo

A escolha de um estilo para a coleta de dados é um dos principais problemas enfrentados pelos pesquisadores da área da prosódia, pois a natureza dos dados pode interferir na análise. Por exemplo, como dito anteriormente, Cunha (2006) cita que o estilo de leitura tem a vantagem de comparabilidade maior entre sujeitos, porém é um estilo de fala pouco natural, enquanto a fala espontânea preserva as características prosódicas, mas não é possível fazer uma comparação mais segmental e más condições de gravação podem influenciar também.

Labov (2008 [1972]) em seu texto “o isolamento de estilos contextuais” discute sobre como diferentes estilos influenciam a fala dos sujeitos, tendo como sua base o nível de atenção prestado à fala. Assim, o autor elabora uma escala da fala em quatro estilos: A. Fala fora da entrevista sociolinguística ou com uma terceira pessoa; B. Fala em situação de entrevista; C. Estilo de leitura e D. Lista de palavras. Sendo “D” a fala mais monitorada e “A” a fala menos monitorada. Ao observar esses estilos se percebe que a leitura está entre os estilos mais monitorados da fala, não cedendo um espaço muito grande à variação. Com base nisso, para este trabalho, fez-se a escolha de analisar apenas o estilo de recontagem de história, pois ao recontar a história, a fala aproxima-se mais de um estilo menos monitorado. Assim, permitindo uma maior liberdade dos parâmetros prosódicos, o que acarreta um dado minimamente próximo da fala natural.

Também tendo estilo como variação entre sujeitos, Labov (1990) elabora em seu trabalho sobre as diferenças entre a fala dos sexos masculino e feminino. O autor aponta para o fato de que os homens tendem a usar com mais frequência as formas não padrão da língua, ficando assim, quase que isentos do estigma social que essas formas carregam, enquanto as mulheres usam com muita frequência as formas padrão, carregando consigo o prestígio atrelado a tais formas. Com base nisto, o autor elabora o princípio de que as mulheres lideram os processos de mudança linguística na língua.

3. Objetivos e hipóteses

O objetivo desta pesquisa é analisar e apontar as possíveis diferenças prosódicas entre algumas variedades do português brasileiro, assim como contribuir com a criação de um *corpus* que possa servir para o estudo da variação dialetal do português brasileiro em seus aspectos fonético-acústicos. Partindo disso, as hipóteses norteadoras do presente estudo são:

- Baseado em concepções impressionísticas, os falantes de Minas Gerais terão uma taxa de elocução e articulação maiores do que as outras variedades;
- Com base na ideia de que as mulheres lideram a mudança linguística (LABOV, 1990), espera-se que haja mais diferenças entre os estados entre as mulheres do que entre os homens;
- Com base nos comentários de Amaral sobre a variedade dita caipira de São Paulo, os falantes de São Paulo terão os menores valores do desvio padrão de f_0 em relação às outras variedades.

4. Metodologia

4.1. *Corpus*

Ao todo, o *corpus* é composto por 43 gravações de 22 sujeitos, dos sexos masculino e feminino, das capitais de sete estados brasileiros: um homem da Bahia, um homem do Espírito Santo, dois homens e duas mulheres de Minas Gerais, uma mulher e um homem do Piauí, três mulheres e dois homens do Rio de Janeiro, uma mulher de Santa Catarina e três homens e cinco mulheres de São Paulo. Os participantes têm entre 20 e 30 anos de idade e foram escolhidos por meio de contatos próximos e redes sociais.

Cada participante leu um trecho do conto "O monge desastrado" e, em seguida, recontou a história com suas próprias palavras. Inicialmente, para uma pesquisa de iniciação científica, os próprios participantes gravaram suas falas utilizando o aplicativo de gravação de voz "Gravador de Voz Fácil". Esse aplicativo permite que as gravações sejam feitas com uma taxa de amostragem de 48.000 KHz e no formato .wav. Após fazerem as gravações, os participantes enviaram os arquivos de áudio por e-mail.

Em uma segunda coleta de dados, realizada para a monografia, os novos participantes gravaram-se utilizando uma plataforma de experimentos criada especificamente para este trabalho. Tendo tudo isso em vista, há duas gravações por pessoa. As gravações da leitura têm, em média, 5 minutos de duração, enquanto as gravações de recontagem têm entre 3 e 4 minutos. Tanto o texto lido pelos participantes quanto às instruções para a gravação dos áudios encontram-se na seção de anexos deste trabalho.

4.2. Novo método de gravação

A plataforma citada funciona de forma totalmente online. Ao acessá-la, o participante é inicialmente introduzido aos termos e condições da pesquisa. Após concordar com esses

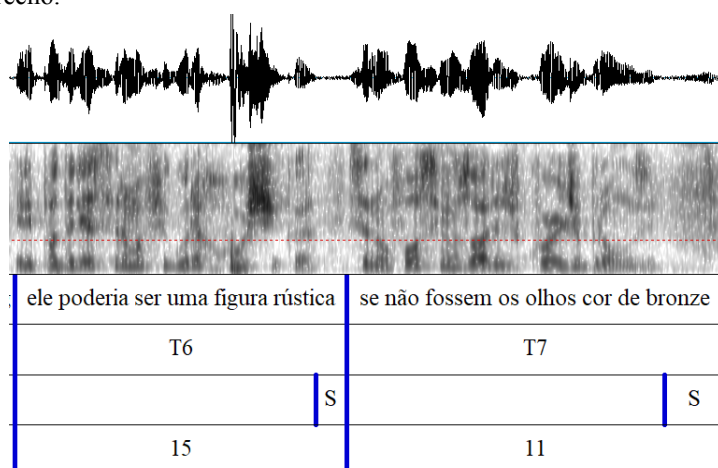
termos, o participante procede ao preenchimento de um formulário contendo informações pessoais, tais como nome, idade, cidade e estado de residência. Após a conclusão do preenchimento, o participante recebe instruções para a gravação, incluindo o trecho do texto "O monge desastrado" e botões para iniciar e encerrar a gravação. Ao terminar a etapa da primeira gravação, uma página semelhante à anterior é exibida, permitindo que o participante realize a gravação da narração do texto lido. Por fim, ao pressionar o botão de enviar, os áudios são armazenados em um servidor privado de acesso apenas dos pesquisadores.

Essa plataforma foi elaborada para simplificar e agilizar o processo de gravação, bem como facilitar o recrutamento de mais participantes. Já que houve uma baixa adesão dos participantes à pesquisa anterior.

4.3. Transcrição dos dados

A segmentação dos áudios foi feita no *software* de acesso livre Praat (BOERSMA, P., WEENINK, D., 2021). Os áudios foram segmentados em trechos de fala separados por uma pausa silenciosa, essas pausas foram marcadas de forma semi-automática com uma ferramenta presente no próprio Praat e os erros foram corrigidos manualmente. Em seguida, realizou-se manualmente a contagem de sílabas fonéticas presentes em cada trecho de fala. O resultado das anotações encontra-se na figura 2:

Figura 2: Segmentação em trechos do texto "O monge desastrado": A parte mais acima é a representação do áudio em ondas sonoras. Abaixo, tem-se, respectivamente, o espectrograma, a transcrição ortográfica, o trecho de fala completo, a marcação de pausas silenciosas e, por fim, a quantidade de sílabas fonológicas presentes no trecho.



Fonte: Autoria própria

4.4. Parâmetros prosódico-acústicos analisados

Após a finalização da segmentação, aplicou-se o *script Prosody Descriptor Extractor* (Barbosa, 2021) que extrai diversos parâmetros prosódico-acústicos dos arquivos de áudio juntamente aos TextGrids, arquivos de base do Praat. Dentre esses parâmetros, são extraídos parâmetros de entonação, intensidade, ritmo e qualidade de voz. Na tabela 1 encontram-se os parâmetros obtidos a partir da extração do *script* e que foram escolhidos para as análises deste trabalho.

Tabela 1: Descrição dos parâmetros prosódicos obtidos pelo *Prosody Descriptor Extractor*

Parâmetro	Descrição	Unidade
Entonação		
f0med	Mediana de f_0	Hz
f0SAQ	Semi-amplitude entre os quartis de f_0	Hz
f0min	Valores mínimos de f_0	Hz
f0max	Valores máximos de f_0	Hz
sdf0peak	Desvio padrão dos picos de f_0	Hz
sdtf0peak	Desvio padrão dos picos de f_0 no eixo temporal	Hz
df0posmean	Primeira derivada da média de f_0 (derivadas positivas)	Hz/quadro
df0negmean	Primeira derivada da média de f_0 (derivadas negativas)	Hz/quadro
df0sdpos	Desvio padrão da primeira derivada de f_0 (derivadas positivas)	Hz/quadro
df0sdneg	Desvio padrão da primeira derivada de f_0 (derivadas negativas)	Hz/quadro
Intensidade		
emph	Ênfase espectral	dB
cvint	Coeficiente de variação da intensidade	%
Ritmo		
sr	Taxa de Elocução	Sílabas/s
ar	Taxa de Articulação	Sílabas/s

Baseado na tabela acima, é preciso trazer uma explicação de alguns dos parâmetros que não são compreendidos de forma intuitiva. O primeiro é a Semi-amplitude entre os quartis de f_0 , ele representa o desvio padrão das medianas de f_0 , ou seja, indica se há variação melódica na voz, quanto maior seus valores mais “vivaz” tende a ser a voz. Os próximos parâmetros são os desvios padrão dos picos de f_0 , que indicam o quanto os valores máximos

de f_0 variam entre si e a distância temporal entre eles, assim indicando se a voz tem a tendência de ser monótona ou não.

Os parâmetros das derivadas de f_0 , tanto a positiva quanto a negativa, indicam se as subidas ou descidas da curva melódica de f_0 ocorrem de forma abrupta ou não, ou seja, se a voz da pessoa repentinamente fica aguda ou grave. Já os desvios padrão desses parâmetros indicam o quão rápido os valores de f_0 aumentam ou diminuem. Por fim, o coeficiente de variação da intensidade, relaciona-se com a variação entre as intensidades das sílabas, contribuindo para uma maior vivacidade e expressividade da voz.

4.5. Considerações da análise

Como o *corpus* aqui analisado é composto por dados de mulheres e homens, é preciso avaliar os sexos separadamente. Já que os valores de f_0 são altamente influenciados por questões anatômicas, principalmente entre mulheres e homens (Beck, 2010). É importante apontar também que alguns estados não tiveram uma quantidade balanceada de participantes, assim, decidiu-se, para este trabalho, analisar apenas os estados da região Sudeste do país (Minas Gerais, Rio de Janeiro e São Paulo), pois foram os estados que mais tiveram participantes. Esse novo recorte totaliza 17 participantes. Um último fator a ser considerado é a semelhança de comportamento entre participantes, ou seja, quando os falantes do mesmo estado apresentam comportamentos semelhantes abre-se a possibilidade de que esse comportamento seja uma característica daquela variedade, pelo contrário tal comportamento pode ser apenas um elemento idiossincrático.

4.6. Análise estatística

Após a extração dos dados numéricos, utilizando o software de programação R³ (R CORE TEAM, 2023), escolheu-se fazer o teste de análise das variâncias (ANOVA). Para tal, fez-se o teste de Shapiro-Wilk para verificar a normalidade das amostras, como nenhuma delas demonstrou ter uma distribuição normal, recorreu-se ao teste não paramétrico de Kruskal-Wallis para avaliar se a hipótese nula é rejeitada ou não. Isto é, se há alguma diferença significativa entre pelo menos dois dos grupos analisados. Assim, se o teste anterior demonstrou que houve diferença, aplicou-se um teste *post hoc* não paramétrico (Teste de Wilcoxon pareado), para fazer comparações entre os estados (SP-RJ, SP-MG, RJ-MG). Para essa análise, adotou-se o nível de significância α de 5% (0.05). Além disso, três variáveis

³ Disponível em: <<http://www.r-project.org>>

independentes foram consideradas: estado e sexo, enquanto as variáveis dependentes foram os próprios parâmetros prosódico-acústicos obtidos por meio do *script*.

Para além da análise anterior, também analisaram-se os dados através de um modelo de regressão de efeitos mistos com os sujeitos sendo a variável aleatória, para verificar se eles apresentariam algum efeito sobre as variedades. No entanto, é importante ressaltar que os resultados para todas as variáveis resposta analisadas apresentaram um efeito menor do que 50%, assim se fez a escolha de trabalhar apenas com a análise ANOVA. Haja vista que “se todas as porcentagens para cada efeito aleatório forem muito pequenas, então os efeitos aleatórios não estão presentes e a modelagem linear de efeitos mistos não é apropriada⁴” (Starkweather, 2010).

5. Discussão dos resultados

5.1. Entonação

A princípio, foi analisada, utilizando o teste de Kruskal-Wallis, a mediana de f_0 . O teste demonstrou que tanto para as mulheres ($\chi^2 = 171.97$, 2, $p < 0.001$), quanto para os homens ($\chi^2 = 53.47$, 2, $p < 0.001$) houve diferença entre os grupos. Assim, verificou-se o teste de Wilcoxon pareado para observar entre quais grupos acontece essa diferença. Para as mulheres, apenas Minas Gerais e o Rio de Janeiro não se diferenciaram entre si. Já para os homens apenas Minas Gerais diferenciou-se das outras variedades.

Tabela 2: Valores da mediana e Semi-amplitude entre os quartis de f_0 (SAQ) em Hertz de mulheres e homens dos estados de MG, RJ e SP.

Estados	Parâmetros	
Mulheres	Mediana	SAQ
Minas Gerais	228 Hz	17.3 Hz
Rio de Janeiro	224 Hz	11.4 Hz
São Paulo	194 Hz	9.68 Hz
Homens		
Minas Gerais	148 Hz	11.0 Hz
Rio de Janeiro	132 Hz	9.69 Hz
São Paulo	131 Hz	9.43 Hz

⁴ If all the percentages for each random effect are very small, then the random effects are not present and linear mixed modeling is not appropriate. (Tradução própria)

Figura 3: Boxplot da mediana de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP.

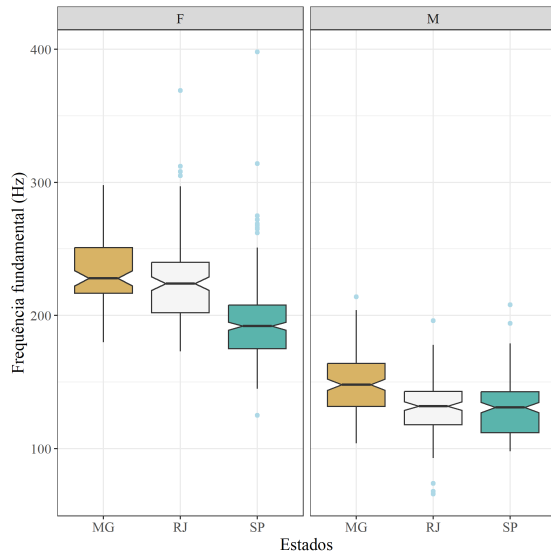
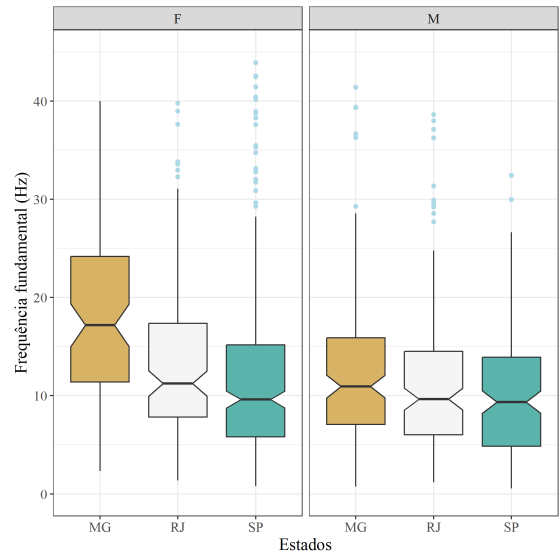


Figura 4: Boxplot da Semi-amplitude entre os quartis de f_0 (SAQ) em Hertz de mulheres e homens dos estados de MG, RJ e SP.



Ao observar a tabela 2, é possível verificar que os valores da mediana de f_0 das mulheres em São Paulo são menores do que dos outros estados, indicando que as vozes das mulheres paulistas tendem a ser mais grave em relação às outras variedades. Enquanto que para os homens os valores de f_0 dos mineiros é mais alto, indicando uma tendência da voz ser mais aguda em relação às outras variedades. (Figura 3)

Ao analisar a Semi-amplitude entre os quartis de f_0 , as mulheres ($\chi^2 = 35.246$, 2, $p < 0.001$) e os homens ($\chi^2 = 6.583$, 2, $p = 0.037$) apresentaram ter uma diferença significativa. No teste *post hoc*, todos os estados se diferenciaram para as mulheres, enquanto os homens só Minas Gerais que se diferenciou dos outros. Observando a tabela 2, percebe-se que os falantes de ambos os sexos de Minas Gerais apresentam os maiores valores, tanto para o valor mínimo quanto para o valor máximo, indicando uma maior variação melódica na fala. (Figura 4)

Tabela 3: Mediana dos valores mínimos e máximos de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP.

Estados	Parâmetros	
Mulheres	Mínimos	Máximos
Minas Gerais	192 Hz	292 Hz
Rio de Janeiro	182 Hz	263 Hz
São Paulo	164 Hz	242 Hz
Homens		
Minas Gerais	115 Hz	190 Hz
Rio de Janeiro	94 Hz	158 Hz
São Paulo	107 Hz	169 Hz

Figura 5: Boxplot dos valores mínimos de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP.

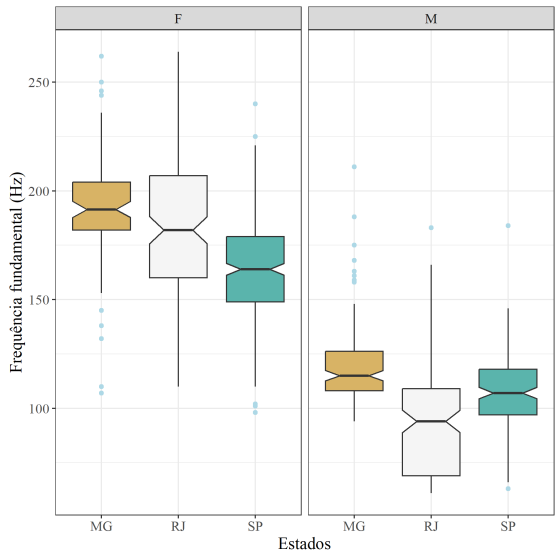
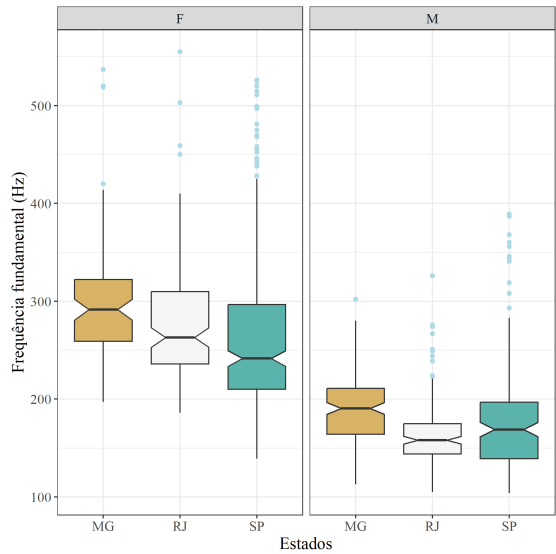


Figura 6: Boxplot dos valores máximos de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP.



Para os valores mínimos de f_0 , as mulheres paulistanas se comportam de forma semelhante ao da mediana de f_0 , ainda apresentando valores menores do que as mulheres das outras variedades (Tabela 3). Neste teste, apenas São Paulo diferenciou-se das outras variedades ($\chi^2 = 93.731$, 2, $p < 0.001$). Já para os homens as três variedades se diferenciaram ($\chi^2 = 36.359$, 2, $p < 0.001$) e também trazendo um resultado semelhante ao da mediana de f_0 , com os mineiros apresentado um valor maior do que as outras variedades, seguido de São Paulo e Rio de Janeiro. (Tabela 3; Figura 5)

Os valores máximos de f_0 também apresentaram resultados semelhantes aos parâmetros de mediana e valores mínimos de f_0 para os estados, todas as variedades se

diferenciam para as mulheres ($\chi^2 = 40.864$, 2, $p < 0.001$) e para os homens apenas Minas Gerais de diferencia das outras variedades ($\chi^2 = 38.384$, 2, $p < 0.001$). Assim, as mulheres de São Paulo ainda apresentam os menores valores dentre as outras variedades. Enquanto para os homens Minas Gerais apresenta valores maiores em relação às outras variedades. (Tabela 3; Figura 6)

O desvio padrão dos picos de f_0 apresentou diferença apenas entre Minas Gerais e o Rio de Janeiro para as mulheres ($\chi^2 = 6.9086$, 2, $p = 0.031$), enquanto para os homens não houve diferença entre as variedades ($\chi^2 = 1.3125$, 2, $p > 0.05$). Com isso em vista, Minas Gerais apresentou os maiores valores para as mulheres (Tabela 4; Figura 7). O que indica também uma variação maior da melodia da voz, indo em convergência aos resultados da Semi-amplitude entre os quartis de f_0 .

Já o desvio padrão dos picos de f_0 no eixo temporal não apresentou diferença significativa entre nenhuma das variedades em ambos os sexos (Para as mulheres: $\chi^2 = 2.8996$, 2, $p > 0.05$; Para os homens $\chi^2 = 7.7186$, 2, $p > 0.05$). (Tabela 4; Figura 8).

Tabela 4: Mediana dos valores de desvio padrão dos picos de f_0 em picos por segundo de mulheres e homens dos estados de MG, RJ e SP.

Estados	Parâmetros	
Mulheres	sdf ₀ peak	sdtf ₀ peak
Minas Gerais	31.8 P/s	0.18 P/s
Rio de Janeiro	18.8 P/s	0.2 P/s
São Paulo	23.4 P/s	0.2 P/s
Homens		
Minas Gerais	21.2 P/s	0.17 P/s
Rio de Janeiro	16.4 P/s	0.24 P/s
São Paulo	20.0 P/s	0.21 P/s

Figura 7: Boxplot do desvio padrão dos picos de f_0 em picos por segundo de mulheres e homens dos estados de MG, RJ e SP.

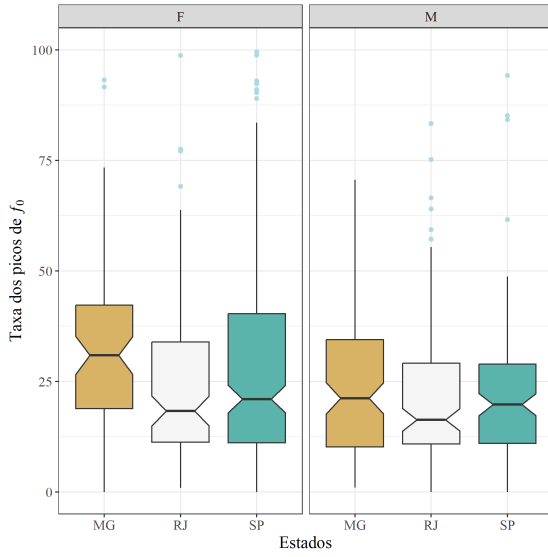
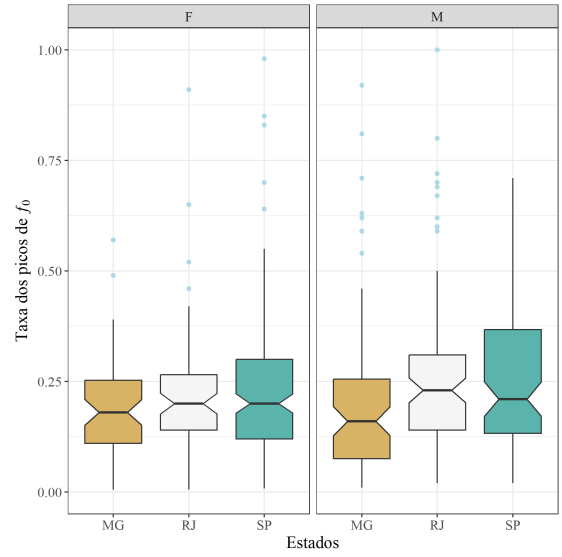


Figura 8: Boxplot do desvio padrão dos picos de f_0 no eixo temporal em picos por segundo de mulheres e homens dos estados de MG, RJ e SP.



Os valores das derivadas positivas de f_0 , não apresentaram diferença significativa entre nenhuma das variedades para as mulheres ($\chi^2 = 0.37971$, 2, $p > 0.05$). Já para os homens, Minas Gerais se diferenciou apenas do Rio de Janeiro ($\chi^2 = 20.369$, 2, $p < 0.001$) e apresentando os maiores valores (Tabela 5; Figura 9). Indicando que os homens mineiros tendem a subir os valores de f_0 com mais frequência.

De forma semelhante ao parâmetro anterior, as mulheres não apresentaram diferenças no desvio padrão das derivadas positivas de f_0 ($\chi^2 = 0.096$, 2, $p > 0.05$). Já os homens apresentaram uma diferença significativa ($\chi^2 = 16.042$, 2, $p < 0.001$) apenas entre Minas Gerais o Rio de Janeiro, também apresentando os maiores valores de desvio padrão (Tabela 5; Figura 10). O que indica que os homens mineiros aumentam os valores de f_0 de forma mais abrupta.

Tabela 5: Mediana e desvio padrão das derivadas positivas de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP.

Estados	Parâmetros	
Mulheres	Mediana	Desvio padrão
Minas Gerais	4.38 Hz	3.44 Hz
Rio de Janeiro	4.68 Hz	3.76 Hz
São Paulo	4.36 Hz	3.56 Hz
Homens		
Minas Gerais	3.53 Hz	2.62 Hz
Rio de Janeiro	2.44 Hz	1.79 Hz
São Paulo	2.9 Hz	2.23 Hz

Figura 9: Boxplot das derivadas positivas de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP.

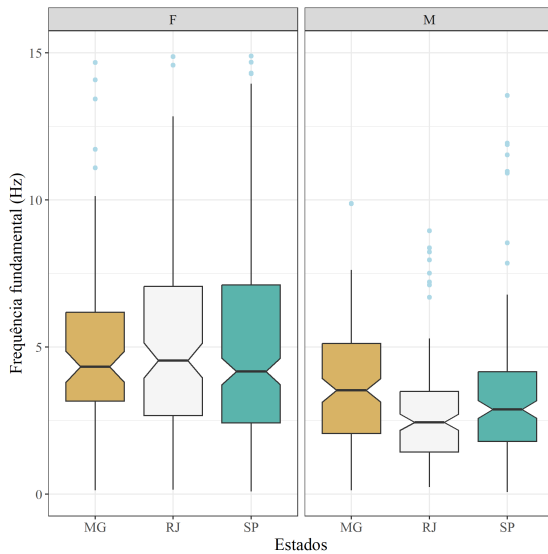
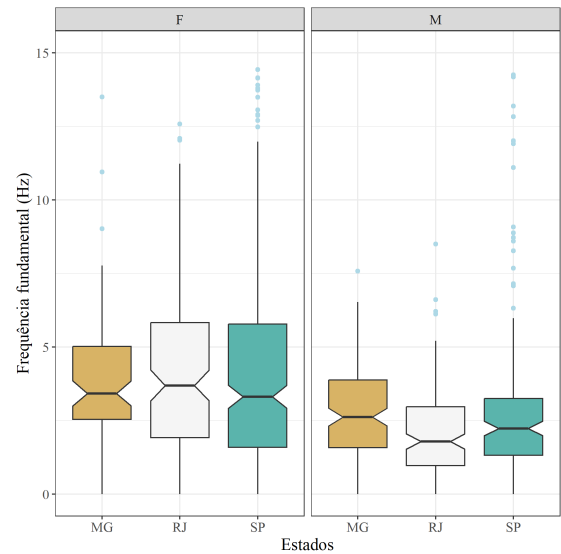


Figura 10: Boxplot do desvio padrão das derivadas positivas de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP.



Analisando as derivadas negativas de f_0 , apenas Minas Gerais se diferenciou das outras variedades para as mulheres ($\chi^2 = 19.779$, 2, $p < 0.001$), enquanto para os homens apenas São Paulo se diferenciou das outras variedades ($\chi^2 = 12.334$, 2, $p < 0.05$). Assim, para as mulheres, Minas Gerais apresentou os menores valores (Tabela 6; Figura 11) seguidos do Rio de Janeiro e São Paulo, ou seja, as descidas das curvas melódicas de f_0 das mulheres mineiras são mais íngremes.

Tabela 6: Mediana e desvio padrão das derivadas negativas de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP.

Estados	Parâmetros	
Mulheres	Mediana	Desvio padrão
Minas Gerais	-3.46 Hz	3.89 Hz
Rio de Janeiro	-2.37 Hz	3.28 Hz
São Paulo	-2.36 Hz	2.71 Hz
Homens		
Minas Gerais	-2.04 Hz	2.40 Hz
Rio de Janeiro	-2.21 Hz	2.64 Hz
São Paulo	-1.50 Hz	1.97 Hz

Figura 11: Boxplot das derivadas negativas de f_0 em Hertz de mulheres e homens dos estados de MG, RJ e SP.

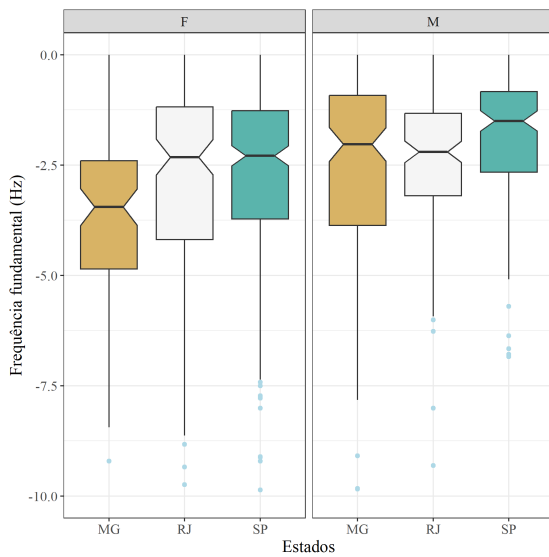
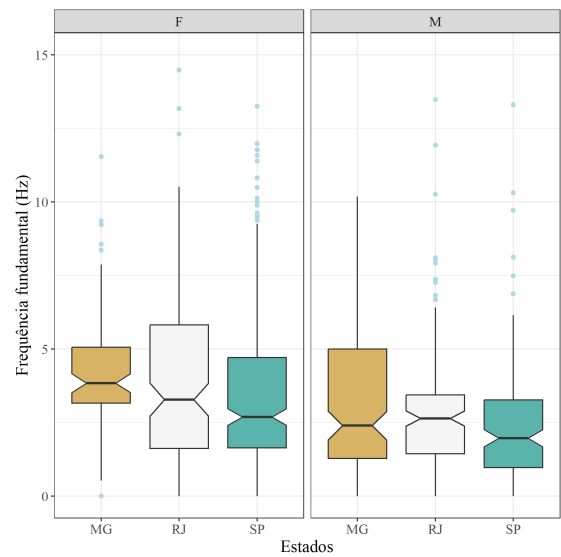


Figura 12: Boxplot do desvio padrão das derivadas negativas de f_0 em Hertz De mulheres e homens dos estados de MG, RJ e SP.



O desvio padrão das derivadas negativas de f_0 apresentou diferença para as mulheres de São Paulo ($\chi^2 = 16.28$, 2, $p < 0.001$), sendo que essa variedade apresentou os menores valores de variação. Já para os homens apenas Minas Gerais e São Paulo se diferenciaram ($\chi^2 = 8.191$, 2, $p < 0.05$), com o último apresentando os menores valores de variação. (Tabela 6; Figura 12)

5.2. Intensidade

Partindo para a análise sobre a intensidade, a ênfase espectral apresentou diferença entre as três variedades para as mulheres ($\chi^2 = 61.334$, 2, $p < 0.001$) e ao observar os valores

da tabela 7 percebe-se que é possível fazer uma escala dos valores, com Minas Gerais apresentando o maior valor seguido do Rio de Janeiro e por fim São Paulo. Para os homens apenas o Rio de Janeiro se diferenciou das outras variedades ($\chi^2 = 108.23$, 2, $p < 0.001$) apresentando os maiores valores (Tabela 7; Figura 13). A partir disso, pode-se inferir que a fala das mulheres mineiras pode ser percebida com mais esforço vocal, enquanto para os homens essa característica se aplica aos cariocas.

Tabela 7: Valores da ênfase espectral (dB) e coeficiente de variação da intensidade em porcentagem de mulheres e homens dos estados de MG, RJ e SP.

Estados	Parâmetros	
Mulheres	Ênfase espectral	Coeficiente de variação
Minas Gerais	4.6 dB	16%
Rio de Janeiro	3.3 dB	17%
São Paulo	2.5 dB	15%
Homens		
Minas Gerais	1.65 dB	18%
Rio de Janeiro	3.3 dB	13%
São Paulo	2.0 dB	16%

Figura 13: Boxplot da ênfase espectral em decibéis para mulheres e homens dos estados de MG, RJ e SP.

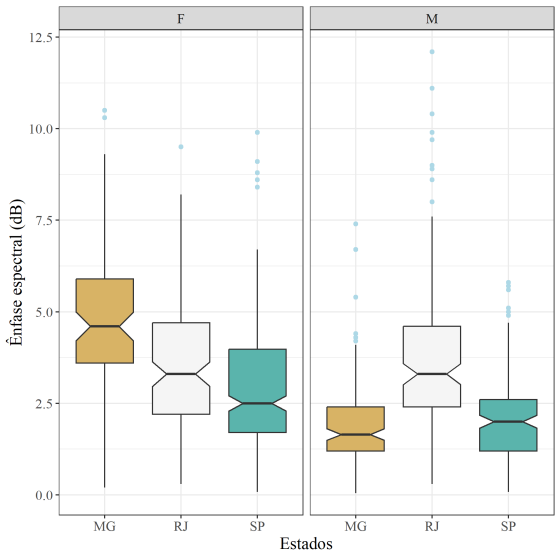
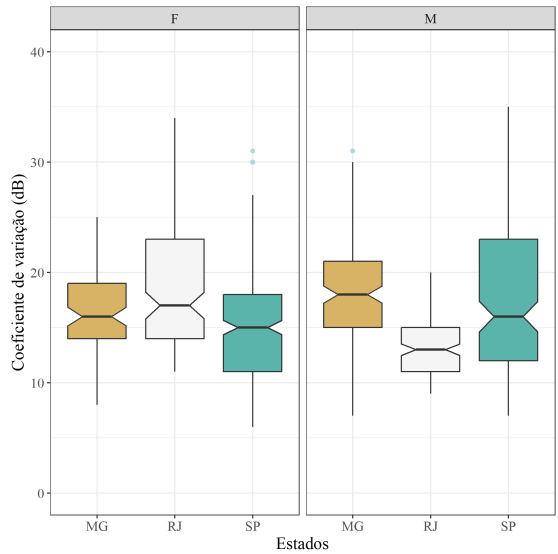


Figura 14: Boxplot do coeficiente de variação da intensidade em decibéis de mulheres e homens dos estados de MG, RJ e SP.



Agora para o coeficiente de variação da intensidade, apenas as mulheres de São Paulo se diferenciaram das outras variedades ($\chi^2 = 34.128$, 2, $p < 0.001$) apresentando o menor

valor, no entanto ainda se mantendo próximo das outras variedades. Para os homens, novamente, apenas o Rio de Janeiro se diferenciou das outras variedades ($\chi^2 = 95.329$, 2, $p < 0.001$) desta vez, com o menor valor dentre as outras variedades. (Tabela 7; Figura 14). Assim, é possível aferir que as mulheres de São Paulo apresentam uma menor variação na intensidade das sílabas em relação às outras variedades, já para os homens essa característica se aplica aos cariocas. Enquanto os maiores valores são apresentados por Minas Gerais.

5.3. Ritmo

Por fim, foram analisadas as taxas de elocução e articulação. Para a taxa de elocução apenas as mulheres de Minas Gerais se diferenciaram das outras variedades ($\chi^2 = 84.697$, 2, $p < 0.001$) e apresentaram os maiores valores (Tabela 8; Figura 15). Para os homens houve diferença apenas entre o Rio de Janeiro e São Paulo ($\chi^2 = 7.3139$, 2, $p < 0.05$) sendo São Paulo a variedade que apresentou os menores valores. Ao observar esse parâmetro percebe-se que as mulheres de Minas Gerais apresentam uma fala considerada como mais “rápida” em comparação às outras variedades. Uma observação interessante de se fazer é que, ao realizar a transcrição das gravações, foi possível perceber quedas silábicas e a coarticulação de segmentos fonéticos na fala dos mineiros, o que contribui para a percepção dessa fala como mais “rápida”.

Como já dito anteriormente na seção 2, as taxas de elocução e articulação são bem semelhantes, apenas se diferenciando na inclusão da duração dos silêncios na conta da taxa de elocução. Assim, seus resultados também são semelhantes. Com isso em mente, para a taxa de articulação, apenas as mulheres de Minas Gerais se diferenciaram das outras variedades ($\chi^2 = 22.228$, 2, $p < 0.001$). Apresentando assim o maior valor da amostra. Para os homens, apenas São Paulo se diferenciou das outras variedades ($\chi^2 = 11.69$, 2, $p < 0.05$), apresentando novamente os menores valores (Tabela 8; Figura 16).

Tabela 8: Medianas das taxas de elocução e articulação em sílabas por segundo de mulheres e homens dos estados de MG, RJ e SP.

Estados	Parâmetros	
Mulheres	Elocução	Articulação
Minas Gerais	5.3 S/s	6.65 S/s
Rio de Janeiro	3.9 S/s	6.0 S/s
São Paulo	3.6 S/s	5.5 S/s
Homens		
Minas Gerais	4.1 S/s	6.1 S/s
Rio de Janeiro	4.2 S/s	5.9 S/s
São Paulo	3.7 S/s	4.95 S/s

Figura 15: Boxplot da taxa de elocução em sílabas por segundo de mulheres e homens dos estados de MG, RJ e SP.

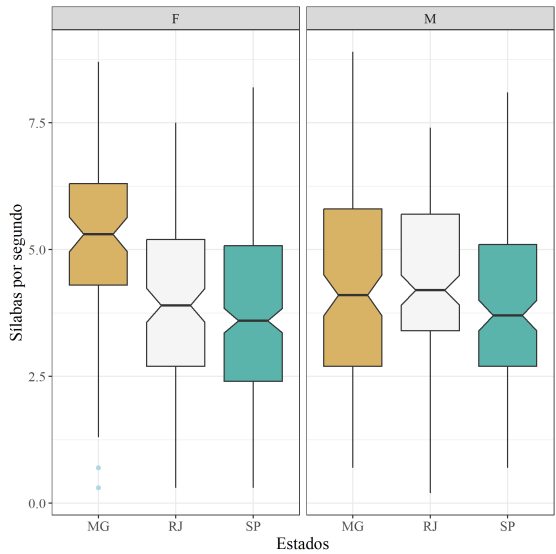
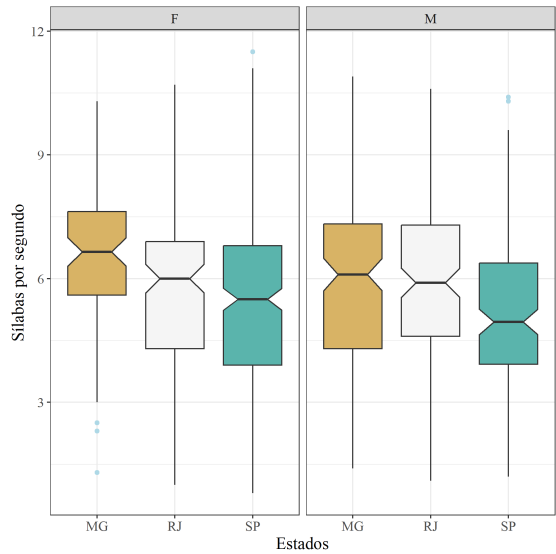


Figura 16: Boxplot da taxa de articulação em sílabas por segundo de mulheres e homens dos estados de MG, RJ e SP.



6. Conclusão

Com base nos parâmetros prosódicos, foi possível distinguir e caracterizar as três variedades analisadas (MG, RJ, SP), evidenciando que a prosódia, em conjunto com aspectos fonológicos, lexicais e sintáticos, é capaz de diferenciar as variedades do português brasileiro.

Em relação aos correlatos acústicos da entoação, os principais resultados se concentraram na mediana de f_0 e seus valores mínimos e máximos. Observou-se que as mulheres paulistanas tendem a ter valores mais baixos de f_0 , em torno de 194 Hz, indicando vozes mais graves em comparação com as outras variedades. Quanto aos desvios padrão tanto da curva melódica quanto dos valores medianos de f_0 , Minas Gerais apresentou os maiores valores, sugerindo uma fala mais expressiva e melódica.

Nos parâmetros de intensidade, as mulheres de Minas Gerais demonstraram um maior esforço vocal, cerca de 4.6 dB, enquanto para os homens essa característica foi observada nos homens do Rio de Janeiro, em torno de 3.3 dB. Em relação ao coeficiente de variação da intensidade, as mulheres paulistanas e os homens cariocas apresentaram os menores valores.

Quanto aos parâmetros de ritmo, ambas as taxas indicaram que as falantes mineiras têm uma fala mais rápida, seguidas pelo Rio de Janeiro e São Paulo, corroborando com a hipótese de que os falantes mineiros apresentariam as maiores taxas de articulação e elocução.

Observou-se também que as variedades puderam ser diferenciadas de maneira mais eficaz com as falantes mulheres do que com os falantes homens, sugerindo um efeito de gênero na diferenciação das variedades estudadas. Com isso, pode-se hipotetizar que homens e mulheres usam diferentes estratégias em sua fala, inclusive na prosódia, para caracterizar sua fala.

Por fim, um resultado interessante de apontar é que a prosódia foi um fator determinante para a diferenciação de Minas Gerais das outras regiões. Com esse resultado em vista, é interessante para pesquisas futuras realizar testes perceptivos com foco na prosódia, para assim observar esse efeito da prosódia sobre essas variáveis e complementar os resultados encontrados aqui.

7. Bibliografia

- AMARAL, A. **O dialeto caipira**. 2ª Edição. Livro Digital nº 306. São Paulo: Organização Simões, 2019. [1920].
- ATKINSON, K. **Language identification from non-segmental cues**. *Working papers in phonetics* (UCLA), 10, p. 85-89, 1968.
- BARBOSA, P. A. **Manual de Prosódia Experimental**. Livro Digital. Campinas, SP: Editora da Abralín, 2022.
- BARBOSA, P. A. **Prosódia**. Parábola Editorial, São Paulo, 2019.
- BARBOSA, P. A. **Prosody Descriptor Extractor**. (Versão 3.0) [Programa de computador]. Disponível em: <<https://github.com/pabarbosa/prosody-scripts/tree/master/ProsodyDescriptorExtractor>>, 2021.
- BECK, J. M. **Organic variation of the vocal apparatus**. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The Handbook of Phonetic Sciences* (pp. 523–602). Wiley-Blackwell. 2010.
- BIADSY, F.; HIRSCHBERG, J. **Using Prosody and Phonotactics in Arabic Dialect Identification**. Interspeech. ISCA. 2009.
- BOERSMA, P.; WEENINK, D. **Praat: doing phonetics by computer**. 2022. Disponível em: <https://www.fon.hum.uva.nl/praat/>. Acesso em: 14 set. 2023.
- CLOPPER, C.G., SMILJANIC, R. **Effects of gender and regional dialect on prosodic patterns in American English**. *Journal of Phonetics*. 2011. April 1; 39(2): 237–245.
- CONSTANTINI, A. C. **Caracterização prosódica de sujeitos de diferentes variedades de fala do Português brasileiro em diferentes relações Sinal-Ruído**. 115f. Tese de Doutorado - Instituto dos Estudos da Linguagem, Unicamp, Campinas, 2014.
- CUNHA, C. **Atlas lingüístico do brasil: uma análise das questões de prosódia**. In: Mota, J.A.; Cardoso, S.A M.(Org.). *Documentos 2: projeto atlas lingüístico do Brasil*. Salvador: Quarteto, 2006.
- ETMAN, A; LOUIS, A. A. **American dialect identification using phonotactic and prosodic features**. SAI intelligent systems conference, 2015.
- ERIKSSON, A. **The disguised voice: imitating accents or speech styles and impersonating individuals**. In: C. Llamas & D. Watt (Eds.), *Language and Identities*. Edinburgh: Edinburgh University Press. 2010. p. 86–96
- FLETCHER, J. **The prosody of speech: Timing and rhythm**. In: W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The Handbook of Phonetic Sciences* (pp. 523–602). Wiley-Blackwell. <https://doi.org/10.1002/9781444317251>. ch15. 2010.
- FREITAG, R. et al. **Como os brasileiros acham que falam? Percepções Sociolinguísticas de universitários do sul e do nordeste**. *Todas as Letras*, São Paulo, v. 18, n. 2, p.64-84, 2016.

LABOV, W. **Padrões sociolinguísticos**. Tradução: Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo, Parábola Editorial, 2008. [1972]. p. 63-138.

LABOV, W. **The intersection of sex and social class in the course of linguistic change**. *Language variation and Change*. 1990; 2(2), p. 205-254. doi:10.1017/S0954394500000338

MORAES, J. A. **Intonation in Brazilian Portuguese**. In: HIRST, D.; DI CRISTO, A. (Ed.) *Intonational Systems: a Survey of Twenty Languages*. Cambridge: MIT Press, 1998, Chap. 10.

NASCENTES, A. **O linguajar carioca**. 2ªed. completamente refund. Rio de Janeiro, RJ: Simões, 1953.

PERES, D. O. **O papel da prosódia na identificação das variedades regionais do português brasileiro**. 152 fls. Dissertação de mestrado - Faculdade de Filosofia, Letras e Ciências Humanas, Programa de Pós-Graduação em Filologia e Língua Portuguesa, Universidade de São Paulo, São Paulo, 2011.

Projeto AliB, disponível em: <<https://alib.ufba.br>> acesso em: 19 out. 2023.

Projeto AMPER-POR, disponível em: <http://www.varialing.eu/?page_id=272> acesso em: 19 out. 2023.

R CORE TEAM. **A Language and Environment for Statistical Computing** (Version 3.6.1). [S.l.: s.n.], 2023.

STARKWEATHER, J. **Linear mixed effects modeling using R**. *Unpublished Manuscript* (2010).

ANEXOS

- **Trecho do texto “O monge desastrado”:**

Manuel desceu a escada, maldizendo a sua vida. Por mais que se esforçasse, estava sempre se metendo em problemas. Tristonho, foi ver o irmão Bernardo de Santa Maria. E este deu um sorriso, com uma ternura bondosa. Muito baixo, gordo, de careca brilhante e duas bochechas encarnadas como maçãs maduras, ele poderia ser uma figura rústica se não fossem os olhos cor de bronze, mais brilhantes do que estrelas. Olhos de um homem sereno, inteligente e feliz.

Bernardo, em vez de fazer perguntas, o mandou trabalhar ao lado do forno onde os irmãos faziam grandes pães redondos. Um calorzinho bom devolveu-lhe as forças. E colocou as mãos à obra sem saber o porquê sentia a cabeça girar e um vazio no estômago. Mas o motivo era bem simples. O pão quente e o leite morno exalavam um cheiro de dar água na boca. Não se atreveu, no entanto, a pedir nada. Afinal de contas, estava de castigo.

Para enganar a fome, enterrou as mãos na massa até ao cotovelo e fez uma grande bola. Depois olhou em volta de soslaio. Talvez não reparassem se ele comesse um bocado de massa crua. Ninguém reparou. Então, em vez de fazer os movimentos certos, começou a amassar ao acaso dobrando e redobrando aquela mistura de farinha e água que iria se transformar em pão. Uma ideia genial passou pela sua cabeça. Se fosse buscar outros mantimentos; Em vez de pão poderia fazer uma bola para rechear de carne. Ou um doce. Assim poderia ir comendo bocadinhos disto e daquilo sem que ninguém percebesse.

Foi dito e feito! Guloso como era, começou pela manteiga, depois foi ao leite, raspas de limão, ovos... O frei Bernardo o olhava discretamente sem dizer nada. Manuel tinha recuperado as cores e girava pela cozinha na maior correria.

Em cima da mesa, em vez de um pão redondo, tinha várias tigelinhas de massa cheias de creme. Meteu tudo no forno e foi se sentar. Estava cansadíssimo! Alguns minutos de descanso não seriam notados. Ele se encostou na parede, fechou os olhos e deixou que uma dormência suave fosse tomando conta do seu corpo e espírito. Ele estava tão bem ali! Flutuando dentro de si mesmo que não ouviu os sinos anunciarem que amanhecia. Nem foi à igreja para rezar a Prima. Escorregou para debaixo da mesa e dormia profundamente deitado no chão.

Quando, algumas horas depois, a voz do prior se fez ouvir na cozinha, Manuel quase desmaiou de susto. Não pensava em nenhuma desculpa possível para se justificar. Levantou-se passado de vergonha, com a roupa cheia de manchas, o cabelo bagunçado e os olhos inchados. E foi naquela triste figura que o prior entrou no refeitório. Ali estava o frei

Diogo, os outros monges e irmãos, todos com ar muito solene. Com certeza iam expulsá-lo do convento. Abaixou a cabeça, esperando ouvir palavras terríveis, e só então reparou que, na mesa estava um tabuleiro cheio de pastéis que alguém polvilhou de canela.

- Este rapaz é um exemplo para todos nós - exclamou o prior, apontando para o tabuleiro com um gesto breve. - Ele fez uma bela ação, merece uma recompensa.

Manuel ainda olhou à procura do tal rapaz que seria recompensado, e ficou perplexo quando percebeu que era ele próprio, porque não se lembrava de ter feito nada excepcional.

- Ele passou a noite trabalhando sem descanso - continuou o prior no mesmo tom de aprovação.

- Com certeza não poupou esforços para inventar os melhores pastéis que já comi na minha vida. Podem até ser vendidos para fora, o que ajudará muito as finanças do mosteiro.

Até frei Diogo acenou que sim. E não é que ele parecia contente? O prior se virou com um sorriso aberto:

- A partir de hoje você será o nosso doceiro chefe! O pobre rapaz ficou sem saber se ria ou se chorava. Nunca lhe tinham feito um elogio! As palavras do prior encheram seu peito com uma alegria nova, desconhecida. Mas o pior é que não se lembrava da receita inventada por acaso na noite anterior. Era preciso confessar a verdade. E coragem?

Foi o irmão Bernardo quem o socorreu. Primeiro, fez sinal para que ficasse quieto. Depois, a sós, explicou que tinha reparado em todas as voltas que ele deu, em tudo o que usou e que juntos fariam uma nova dose nesse mesmo dia.

- Fui eu quem colocou a canela - confessou. - E ficaram tão deliciosos que só por minha conta já comi seis! Você será famoso, Manuel. Você e os seus pastéis de Belém!

● **Instruções para a entrega das gravações:**

Só lembrando que será feita a gravação da leitura e da recontagem do conto com suas próprias palavras.

Para fazer as gravações é preciso instalar o aplicativo gratuito "gravador de voz fácil", ir em sua configuração, colocar o formato de arquivo em .wav (PCM) e a taxa de amostragem em 48kHz.

Após feitas as duas gravações, elas precisam ser enviadas ao e-mail 4015luis.henrique@gmail.com com o assunto sendo "pesquisa [complete com seu estado]" e no mesmo e-mail enviar uma confirmação de aceite do TCLE no seguinte formato:

Eu___ li o TCLE da resolução 510/2016 e concordo com seus termos.

- **Instruções do uso do aplicativo “Gravador de voz fácil”:**

ANDROID

- Instale o aplicativo “gravador de voz fácil” e o abra.
- Para configurar o aplicativo aperte nos três pontos no canto superior direito e entre em “configurações”;
- Em seguida vá em afinação;
 - Ao entrar em afinação há uma opção para formato, escolha o formato PCM (.wav);
 - Há também em afinação a opção taxa de amostragem, escolha a opção 48kHz (mais alta);
- Quando terminar, volte para tela inicial do aplicativo e faça a gravação;

IOS

- Instale o aplicativo “gravador de voz fácil” e o abra.
- Ao abrir o aplicativo, no canto inferior direito há uma engrenagem de “configurações”.
- Ao abrir as configurações aperte na opção de “formato”.
 - Em formato do arquivo escolha o .wav (PCM).
 - Em taxa de amostragem escolha a opção de 48kHz (mais alta).
- Quando terminar, volte para tela inicial do aplicativo e faça a gravação;

INSTRUÇÕES PARA A GRAVAÇÃO

- Grave sua voz em ambiente silencioso.
- Sugiro que o microfone do celular seja posicionado a 10 cm da boca (tudo bem se não forem 10cm exatos).
- Não é preciso gritar ou falar mais alto, use seu tom de voz normal (como se estivesse conversando).