

RELATÓRIO CIENTÍFICO

A ACOMODAÇÃO DIALETAL E A ESTABILIDADE DE  
PADRÕES SOCIOLINGUÍSTICOS NA FALA ADULTA

LIVIA OUSHIRO

Relatório Científico apresentado ao Programa  
de Pós-Graduação em Linguística da UFRJ,  
referente à pesquisa financiada pelo Programa  
Nacional de Pós-Doutorado da CAPES.

Supervisora: Prof<sup>a</sup> Dr<sup>a</sup> Christina Abreu Gomes  
Linha de Pesquisa: Variação e Mudança Linguística

Rio de Janeiro  
30 de abril de 2016

# 1 Introdução

Este relatório compreende as atividades realizadas no período de junho de 2015 a fevereiro de 2016, referente ao desenvolvimento da pesquisa “A acomodação dialetal e a estabilidade de padrões sociolinguísticos na fala adulta”, financiada por bolsa do Programa Nacional de Pós-Doutorado da CAPES. O cancelamento da bolsa foi realizado devido à convocação da pós-doutoranda para assumir o cargo de Professora Doutora na Universidade Estadual de Campinas, em Regime de Dedicção Integral à Docência e à Pesquisa, após ter sido aprovada em concurso público realizado em setembro de 2015.

O projeto propôs a análise da pronúncia variável de vogais médias pretônicas, em palavras como *relógio* e *roseira*, na fala de migrantes paraibanos residentes nas cidades do Rio de Janeiro e de São Paulo, com vistas a analisar os mecanismos que presidem processos de acomodação dialetal, tópico ainda relativamente pouco estudado. De modo amplo, esse objeto concerne à teoria linguística no que diz respeito à estabilidade do sistema linguístico na fala adulta, ao mesmo tempo que contribui para um quadro mais completo da Teoria da Variação e da Mudança (Weinreich *et al.*, 2006 [1968]), ao descrever e analisar a fala de indivíduos que atualmente constituem grande proporção das populações que hoje habitam as grandes metrópoles.

A acomodação dialetal por parte de migrantes apenas recentemente vem recebendo a atenção de estudos sociolinguísticos no país e ainda carece de análises mais aprofundadas que permitam sistematizações e generalizações. Para tal empreendimento, comunidades urbanas como Rio de Janeiro e São Paulo

se apresentam como locais privilegiados, devido à ampla heterogeneidade sociodemográfica de sua população. Embora os fluxos migratórios no Brasil tenham presenciado maior polarização para diversas regiões nos últimos anos, os estados do Nordeste continuam sendo aqueles com maior média de evasão populacional e os estados do Sudeste aqueles que mais recebem migrantes, com destino sobretudo para essas duas capitais (Oliveira *et al.*, 2011).

Nesse cenário, o presente projeto teve os objetivos de: (i) tratar da estabilidade da fala adulta e dos efeitos de diferentes variáveis independentes no processo de acomodação dialetal; (ii) disseminar novos métodos e ferramentas de análise sociolinguística; e (iii) oferecer uma descrição mais ampla do português falado em comunidades urbanas.

No que segue, reportam-se os passos tomados na presente pesquisa para atingir esses objetivos. Após retomar o cronograma de atividades proposto no projeto e apresentar uma breve revisão bibliográfica dos estudos pertinentes a esta pesquisa, descrevem-se a organização da amostra analisada (seção 2.2), o desenvolvimento e uso de novas ferramentas de análises na plataforma R (R Core Team, 2015) (seção 2.3), a análise dos dados (seção 2.4) e os passos futuros que se projetam para a continuidade da pesquisa (seção 2.5). A seção 3 relata outras atividades realizadas durante o período de usufruto da bolsa – como participações em encontros científicos, publicações e minicursos ministrados – que, embora não necessariamente relacionadas diretamente ao presente tópico de estudo, contribuíram para a disseminação de métodos e para o intercâmbio com outros pesquisadores com interesses afins, tarefas que também constituíam os objetivos deste projeto.

## 2 Atividades da pesquisa

O Quadro 1 apresenta a lista de atividades realizadas durante o período de vigência da bolsa.

Lista de Atividades	2015						2016		
	J	J	A	S	O	N	D	J	F
<b>Estudo da bibliografia</b>									
Acomodação de fala	✓	✓							
Vogais médias pré-tônicas		✓	✓						
<b>Amostras</b>									
Coleta da amostra PB-RJ	✓	✓	✓						
Organização das amostras			✓	✓					
Novas ferramentas	✓	✓	✓	✓	✓	✓			
<b>Análises</b>									
Amostra PB-SP			✓	✓	✓				
Amostra PB-RJ					✓	✓			
Amostra SP2010						✓	✓		
Amostra Censo00							✓	✓	
Análises quantitativas					✓	✓	✓	✓	✓
<b>Disseminação e avaliação</b>									
Participação em congressos		✓	✓		✓	✓			
Curso sobre a plataforma R					✓	✓			
Avaliação de métodos								✓	✓
Redação de artigos científicos									

**Quadro 1:** Cronograma de Atividades

Os meses iniciais (junho a agosto de 2015) foram dedicados ao estudo da bibliografia levantada a respeito de processos de acomodação dialetal na fala de migrantes e de vogais médias pretônicas; neste mesmo período, também se procedeu à coleta da amostra com falantes paraibanos que residem no Rio de Janeiro e à organização dos conjuntos de gravações (ver seção 2.2).

No período de junho/2015 a novembro/2015, desenvolveram-se métodos

para a otimização da análise de vogais médias pretônicas, concomitantemente à sua aplicação nas quatro amostras analisadas (ver seções 2.3 e 2.4). Paralelamente a essas atividades, a bolsista participou de congressos, publicou artigos referentes aos resultados de sua Tese de Doutorado (ver seção 3), e ofereceu um minicurso na UFRJ sobre a plataforma R. O minicurso foi realizado em seis encontros durante os meses de outubro e novembro de 2015 e teve a presença de aproximadamente 25 participantes, dentre os quais se incluem docentes, pós-graduandos e graduandos da Faculdade de Letras.

Ao longo de todo o período, a bolsista também manteve contato regular com sua supervisora, Prof<sup>a</sup> Dr<sup>a</sup> Christina Abreu Gomes, discutindo leituras e os resultados parciais da pesquisa, e participou de reuniões do grupo de estudos por ela coordenado e de reuniões do PEUL. Adicionalmente, estabeleceu e manteve contato com diversos docentes da Faculdade de Letras da UFRJ.

## **2.1 Revisão bibliográfica**

### **2.1.1 Estudos sobre acomodação dialetal**

Os estudos sobre acomodação dialetal no Brasil, que vêm sendo desenvolvidos principalmente na última década, já permitem nomear uma série de variáveis que possivelmente se correlacionam com padrões de adoção de novas variantes linguísticas.

Um dos estudos pioneiros sobre o comportamento linguístico de migrantes é a pesquisa de Alves (1979), que se debruçou sobre as atitudes de 116 nordestinos (pernambucanos e baianos) residentes na cidade de São Paulo,

com relação às variedades nativas e paulistas. A partir de uma amostra estratificada em *nível social* (alto ou baixo), *procedência* (PE ou BA), *proveniência* (capital ou interior) e *tempo de estada em São Paulo* (recém-chegados ou residentes há mais de dois anos), a autora analisou qualitativa e quantitativamente as respostas dos indivíduos a um questionário previamente elaborado, por meio do qual constatou que os falantes de nível socioeconômico baixo têm atitudes positivas quanto às variedades paulistas, ao passo que os falantes de nível alto tendem a atitudes mais favoráveis quanto às variedades nativas. Além disso, Alves (1979) aventou a hipótese de que os falantes de nível baixo assumiriam traços linguísticos de São Paulo como forma de “camuflar-se linguisticamente”, enquanto os falantes de nível alto tenderiam a manter os traços linguísticos nativos; tal questão, contudo, ficou reservada para estudos futuros.

Desde então, alguns estudos passaram a analisar não só as atitudes, mas também a produção linguística de migrantes. A questão da valorização do dialeto de origem ou da nova localidade é indiretamente abordada por Chacon (2012) e Hora & Wetzels (2010, 2011) através da variável *estilo* (no sentido laboviano de “grau de atenção à fala”; ver Labov 2001a). Em sua análise sobre a palatalização de /s/ em coda medial na fala de 10 paulistas residentes em João Pessoa, Chacon (2012) verificou que o estilo de leitura favorece o emprego da realização alveolar típica do estado de São Paulo, fato que a autora atribui ao maior prestígio do dialeto nativo dos falantes em relação à variedade pessoense. De modo semelhante, em uma análise qualitativa da realização de /r/ em coda na fala de 4 paraibanos residentes em São Paulo, Hora & Wetzels (2010, 2011) previram a preferência pelo tepe em estilo de

leitura, como modo de se aproximar da variante de prestígio local; no entanto, constataram que tal estilo parece favorecer a manutenção da forma vernacular fricativa, de modo que os autores concluem que a hipótese de monitoração não se confirma.

Bortoni-Ricardo (1985), por sua vez, analisou a fala de 33 mineiros estabelecidos em Brasília, com ênfase no papel das *redes sociais* para explicar o desfavorecimento de traços linguísticos tipicamente “rurais” ou “caipiras” (a vocalização da lateral palatal /ʎ/, monotongação de ditongos e a marca zero de concordância verbal): aqueles com maior índice de “integração” e “urbanização” apresentaram maior tendência a empregar formas consideradas padrão.

Bortoni-Ricardo (1985) também verificou, quanto ao *sexo/gênero*, o favorecimento das formas padrão pelos falantes de sexo masculino, algo que a autora atribui a pressões do mercado de trabalho e à conseqüente rede social dos migrantes nas grandes cidades: enquanto os homens geralmente exercem atividades profissionais fora do bairro em que residem e têm maior oportunidade de conviver com os moradores nativos, as mulheres costumam exercer atividades domésticas (como donas de casa ou faxineiras) e manter maior contato com parentes e vizinhos. Todavia, a manutenção das formas vernaculares pelas mulheres não é uma constante. Cardoso (2009), em estudo sobre a alternância entre a forma imperativa associada ao indicativo (*leva, traz, vem*) ou ao subjuntivo (*leve, traga, venha*) em uma amostra de 11 cearenses (e seus descendentes) que vivem em Brasília, constatou o desfavorecimento da forma associada ao subjuntivo, predominante nas localidades nordestinas, por seus informantes do sexo feminino.

Alguns trabalhos assinalam o papel da *idade à época da migração* para a assimilação de novos traços linguísticos (p.ex., Bortoni-Ricardo 1985; Soares 2009) – quanto mais jovem, maior é a tendência de emprego de formas da nova localidade. Entretanto, a variável mais analisada em estudos sobre a acomodação dialetal é o *tempo de permanência*; constataram correlações significativas as pesquisas de Leite (2004) (sobre a pronúncia de /-r/ em coda por 8 estudantes rio-pretenses em Campinas), Marques (2006) (sobre a realização de vogais médias pretônicas na fala de 21 paraibanos no Rio de Janeiro), Martins (2008) (sobre a realização alveolar ou palatalizada de /ti/ e /di/ por 7 paraibanos no Rio de Janeiro), Lima & Lucena (2013) (sobre a pronúncia alveolar ou palatalizada de /s/ em coda na fala de 7 paraibanos em Recife), Bieler da Silva (2015) (sobre a pronúncia de /r/ em coda como retroflexo, fricativa ou tepe na fala de nativos de Itanhandu, comunidade no sul de Minas Gerais), e o estudo de Chacon (2012) mencionado anteriormente.

Por outro lado, Cardoso (2009) afirma não verificar correlação evidente entre o tempo de residência em Brasília e os percentuais de emprego da forma imperativa associada ao indicativo; enquanto dois falantes (Viv e Neu) que ali vivem há 15 e 13 anos respectivamente apresentam forte tendência ao emprego dessa forma verbal, outro falante que vive na cidade há 24 anos (Jes) a desfavorece fortemente. Em vez disso, a autora considera que *o grau de identificação* com a capital federal é preponderante para a preferência pela forma imperativa associada ao indicativo.

De fato, questões de identidade são frequentemente invocadas para explicar o comportamento linguístico dos migrantes. Entretanto, tais considerações não raro são traçadas a partir de análises qualitativas de trechos de entrevistas,



como uma explicação *ad hoc* para a manutenção de traços vernaculares por parte de falantes que migraram há muitos anos. Ao lado do trabalho de Cardoso (2009), um dos poucos estudos que analisa a questão *identidade vs. tempo de permanência* de modo mais sistemático é o de Bieler da Silva (2015), que contrasta os resultados de correlação com as variáveis *Identificação com Itanhandu* e *Tempo Fora da Cidade*. Partindo da hipótese de que as pronúncias não prototípicas de /r/ (tepes e fricativas) por parte dos falantes da comunidade sul-mineira estariam relacionadas a graus de identificação com a cidade, a autora estratificou sua amostra de 36 itanhanduenses entre aqueles que se identificam com a cidade e dela não pretendem sair (ou, tendo saído, pretendem voltar), e aqueles que não se identificam com a cidade e pretendem se mudar dela (ou, tendo saído, não pretendem voltar). Bieler da Silva (2015) constatou uma correlação significativa, mas fraca, com a variável *Identificação com Itanhandu*; por outro lado, diferentemente de Cardoso (2009), verificou uma forte correlação com o tempo de permanência fora da cidade, calculado como uma proporção entre os anos de vida em que o falante morou em outra cidade e sua idade. Tais resultados conduzem-na à conclusão de que as pronúncias não prototípicas de /r/ estão mais relacionadas com a exposição a outros dialetos do que com a identidade ou as atitudes manifestadas pelos informantes.

Juntamente às variáveis sociais mencionadas – classe social, estilo, sexo/gênero, tempo de permanência, identificação –, deve-se ressaltar o amplo espectro de variação individual que normalmente se verifica na fala de migrantes. Mendes (2011), por exemplo, em análise univariada sobre a pronúncia de /-r/ na fala de paraibanos em São Paulo, reporta que as taxas de manutenção da

variante aspirada – em média, 75% – variam individualmente de 56% (peso relativo .47) a 95% (P.R. .91). No estudo de Martins (2008), a média de 64% de palatalização de /ti/ e /di/ na fala de seus 7 informantes paraibanos no Rio de Janeiro inclui taxas que vão de 33% (P.R. .10) à acomodação quase categórica (98%, P.R. .95). Além desses autores, Bieler da Silva (2015) documenta tanto a manutenção categórica de retroflexos quanto o emprego quase categórico de fricativas (94%) na fala de diferentes informantes que emigraram de Itanhandu.

Tal variação certamente decorre da longa lista de variáveis que podem atuar no processo de acomodação dialetal; no entanto, o controle de todas elas através dos métodos tradicionais de estratificação de amostras em estudos sociolinguísticos torna-se problemático.

Desse modo, a consideração de múltiplas variáveis sociais, por um lado, é algo que parece imprescindível no estudo da fala de migrantes; no entanto, isso traz consigo a dificuldade em determinar se os padrões de variação observados se devem de fato às variáveis incluídas na análise estatística ou se podem ser mais propriamente atribuídos à contribuição de falantes individuais específicos (seja pela atuação de variáveis não controladas na estratificação da amostra, seja por idiosincrasias).

A ausência de consenso entre essas conclusões sinaliza a importância de novos estudos que se dediquem a testar tais questões sistematicamente, como se começa a fazer na presente pesquisa.

### 2.1.2 A estabilidade linguística da fala adulta

Ainda que o construto teórico de *mudança em tempo aparente* (Labov, 2001b) tenha sido crucial no estudo da variação e da mudança linguística, através da comparação de padrões de usos variáveis entre falantes de diferentes faixas etárias, uma série de estudos tem começado a investigar mais detalhadamente o pressuposto de que a fala adulta se estabiliza por volta da puberdade e reflete padrões linguísticos passados (Lenneberg 1967, *apud* Nycz 2011, Sankoff & Blondeau 2007).

Nesse sentido, Harrington e colegas, em um conjunto de artigos (Harrington *et al.*, 2000a,b, 2005; Harrington, 2006), mostram as mudanças no sistema vocálico na fala de um indivíduo – a rainha da Inglaterra – entre as décadas de 1950 e 1980, ao comparar as medidas de F1 e F2 de vogais monotongais (Harrington *et al.*, 2000a,b) e ditongais (Harrington *et al.*, 2005) em seu pronunciamento anual de Natal aos britânicos. De modo geral, seus resultados mostram que as vogais da rainha, típicas do RP (*received pronunciation*), deslocaram-se durante esse período na direção do inglês padrão do sul (comparando-as com a fala de quatro jornalistas do sexo feminino da BBC, na década de 1980), embora a diferença entre as variedades ainda seja estatisticamente significativa. Ao mesmo tempo, traços salientes e estereotípicos da fala da realeza (como as realizações dos ditongos /aɪ/ e /aʊ/) são mais resistentes à mudança.

Sancier & Fowler (1997), por sua vez, mostram que deslocamentos na pronúncia de um indivíduo são sensíveis ao tempo de exposição a um ambiente linguístico, mesmo quando se trata de línguas distintas, como é o caso do

sujeito de seu estudo: uma estudante brasileira que alterna períodos de vivência nos Estados Unidos (em que faz seu curso de pós-graduação) e no Brasil (a que volta em seus meses de férias). A pesquisa compara a produção do VOT (*voice-onset time*) de oclusivas surdas como /t/ e /p/, aspiradas no inglês e não aspiradas em português, em um momento em que acabou de voltar aos EUA (tendo passado meses no Brasil falando majoritariamente em português) e em um momento em que está prestes a voltar ao Brasil (tendo passado meses nos EUA falando majoritariamente em inglês). As diferenças nas medidas de VOT se mostraram significativas não só em termos de produção, mas também na percepção de nativos brasileiros, que discriminaram graus de “sotaque” entre pares da mesma palavra realizada nos dois momentos.

Estudos sociolinguísticos de painel, que se baseiam em amostras coletadas dos mesmos indivíduos em diferentes momentos, apresentam quadros distintos (mas não contraditórios) sobre a estabilidade de padrões ao longo do tempo. Paiva & Duarte (2003:28), ao resumir e avaliar os resultados das pesquisas que comparam padrões de variação nas amostras Censo 1980 e Censo 2000 do PEUL, constatam grande irregularidade no comportamento dos indivíduos; enquanto alguns demonstram clara estabilidade do sistema linguístico nesse período de 20 anos, outros apresentam mudanças de natureza quantitativa e qualitativa, nem sempre sendo possível identificar qual ou quais variáveis foram responsáveis pela alteração (seja a própria idade, sejam outros fatores como o aumento do nível de escolarização ou o ingresso no mercado de trabalho). De modo semelhante, Sankoff & Blondeau (2007), em sua análise da pronúncia de /r/ como apical [r] ou dorsal [R] no francês de 22 habitantes

de Montréal no início das décadas de 1970 e de 1980, constataram mudanças substanciais na fala de 9 de seus informantes, embora também tenham verificado estabilidade na fala de outros.

Assim, ainda que haja fortes evidências contra a hipótese de Lenneberg sobre a estabilização do sistema linguístico após o período crítico de aquisição, ainda não se sabem quais fatores levam à adoção ou à rejeição de novos padrões por parte de um indivíduo adulto e quais novos padrões são incorporados.

### 2.1.3 A pronúncia de vogais médias pretônicas

A pronúncia das vogais médias pretônicas foi escolhida para esta pesquisa pelo fato de ser diferenciadora de dialetos; os fenômenos de abaixamento ( $e \rightarrow \varepsilon$ ;  $o \rightarrow \text{ɔ}$ ) e alçamento ( $e \rightarrow i$ ;  $o \rightarrow u$ ) são utilizados por Nascentes (1953 [1922]) como critério para sua divisão dialetal entre “falares do Norte” e “falares do Sul” do Brasil.

Já são inúmeros os trabalhos que se dedicaram à análise dessa variável em diversas comunidades brasileiras, de Norte a Sul do país (ver, p.ex., Bisol 1981, Silva 1989, Callou *et al.* 1991, Battisti 1993, Yacovenco 1993, Celia 2004, Marques 2006, Cruz 2012, Pereira 2010, Carmo 2013, *inter alia*<sup>1</sup>). Sua grande maioria se debruça sobre o fenômeno do alçamento, mas alguns deles tratam de ambos os processos, como é o caso das pesquisas de Pereira (2010) – na fala de moradores de João Pessoa, a partir de dados do Projeto Variação Linguística no Estado da Paraíba (ValPB) –, Yacovenco (1993) – sobre a fala

---

<sup>1</sup>Uma síntese das dezenas de dissertações e teses realizadas sobre o tema se encontra em Chaves (2013).

culta carioca –, e Marques (2006) – sobre migrantes paraibanos no Rio de Janeiro. Esses três trabalhos são de especial interesse a esta pesquisa, uma vez que versam sobre os padrões variáveis de vogais médias pretônicas nas comunidades aqui estudadas.<sup>2</sup> Nas breves resenhas a seguir, o foco é sobre os resultados a respeito do processo de abaixamento, para que possam ser contrastados com os resultados obtidos até o momento da presente pesquisa.<sup>3</sup>

Em seu estudo sobre a fala de paraibanos em João Pessoa, Pereira (2010) analisou dados de vogais pretônicas em sílabas CV ou CVC, excluindo casos de hiatos (como em *realmente*), prefixos (como em *refazer*), além de ocorrências em substantivos próprios e em siglas. Em um total de 8.679 dados de /e/ e 6.401 dados de /o/, extraídos de um *corpus* de 60 informantes estratificados em sexo/gênero, três faixas etárias e cinco níveis de escolaridade, a autora contrastou as realizações da variável dependente de modo ternário (i.e., [ɛ, e, i] e [ɔ, o, u]) em relação às seguintes variáveis independentes: Vogal da Sílabla Seguinte, Distância da Sílabla Tônica (em número de sílabas), Classificação Morfológica (substantivo, verbo etc.), Natureza da Pretônica (átona permanente ou átona casual), Contexto Fônico Precedente e Contexto Fônico Seguinte (as duas últimas, codificadas de acordo com ponto de articulação), além das variáveis sociais estratificadoras da amostra. A partir

---

<sup>2</sup>Na comunidade paulistana, há notícia de apenas um trabalho, Zani (2009), que analisa o alçamento de vogais média-baixas para média-altas na derivação *caf[ɛ] → caf[e]teria*, mas não em *caf[ɛ] → caf[ɛ]zinho*. Não se tem notícia de estudos prévios que tratem do alçamento /e/ → [i] ou /o/ → [u], tampouco do abaixamento para [ɛ] e [ɔ], nessa comunidade.

<sup>3</sup>Interessa também a este trabalho o processo de alçamento das vogais pretônicas; no entanto, como esse fenômeno também ocorre amplamente na variedade paraibana, decidiu-se iniciar a análise pelos contextos mais prováveis para a ocorrência de abaixamento – especificamente, aumento na medida de F1 –, que diferencia mais propriamente as variedades do Sudeste em relação às do Nordeste. Ver seção 2.4.1.

da análise realizada com o pacote de programas VARBRUL (na versão de Pintzuk, 1988, *apud* Pereira 2010), a autora constatou que não havia correlações com as variáveis sociais e que apenas três variáveis linguísticas se mostraram pertinentes aos processos de alçamento e de abaixamento. A principal delas é Vogal da Sílabla Seguinte, que sempre foi a primeira a ser selecionada pelo VARBRUL; o abaixamento da vogal [+anterior] foi favorecido pela presença das vogais /u, ε, ɔ, a, ã, ã, õ/ na sílabla subsequente, e o abaixamento da vogal [-anterior] foi favorecido pela presença das mesmas vogais, além da vogal /ũ/. As outras duas variáveis linguísticas se referem ao contexto fonológico da vogal pretônica: /e/ tende ao abaixamento quando precedido e seguido por vibrante posterior (como em *região* e *pergunta*), enquanto /o/ [-alto] tende a ocorrer quando seguido de sibilantes (como em *social*) ou precedido de segmentos alveolares (como em *novembro*), palatais (como em *jogávamos*) e velares (como em *colega*).

Yacovenco (1993), por sua vez, na análise de 3.563 dados de fala de 18 cariocas da Amostra NURC-RJ (estratificados em sexo/gênero, três faixas etárias e três regiões da cidade – Zona Norte, Zona Sul e zona suburbana), constatou que o fenômeno de abaixamento é pouquíssimo frequente: apenas 99 ocorrências (2,8% do total e 10% quando se excluem os dados de alçamento) correspondiam a essa variante.<sup>4</sup> Pelo pequeno número de dados das variantes média-baixas, sua análise binária de regressão logística no VARBRUL (versão 2S), que contrastou [e] vs. [ε], por um lado, e [o] vs. [ɔ], por outro, incluiu

---

<sup>4</sup>Ainda que se possa questionar a comparabilidade desse estudo com aquele de Pereira (2010), cujas amostras se baseiam em diferentes perfis sociais, Marques (2006) sustenta que tais resultados são comparáveis, pois o levantamento de Vianna da Silva (1995, *apud* Marques 2006:87) sobre as pretônicas nos dialetos populares fluminenses verificou percentuais semelhantes aos da fala culta carioca.

apenas as variáveis sociais devido ao grande número de células vazias para as variáveis linguísticas. Os três fatores externos se mostraram correlacionados com a altura das vogais: o abaixamento de /e/ foi favorecido pelos falantes jovens e velhos da amostra (o que aponta para um caso de variação estável), pelos homens e pelos habitantes da Zona Sul (que, seguindo previsão da autora, mostraram-se mais inovadores do que os falantes da Zona Norte); o abaixamento de /o/, de modo semelhante, também foi favorecido por falantes jovens e velhos, do sexo masculino e habitantes da Zona Sul, além das zonas suburbanas.

Por fim, Marques (2006) analisou a fala de 21 informantes paraibanos do sexo masculino que migraram após os 15 anos de idade ao Rio de Janeiro. Seus informantes têm variação quanto à faixa etária, grau de escolaridade e tempo de residência, embora essas variáveis não sejam ortogonais entre si (p.ex., não há informantes entre 15 e 25 anos com nível superior de escolaridade, ou falantes com mais de 50 anos que habitem o Rio de Janeiro há menos de 5 anos).

A Tabela 1, adaptada de Marques (2006:86), compara as proporções das variantes média-baixa, média-alta e alta, para a variável /e/ e para a variável /o/, de acordo com os levantamentos de Yacovenco (1993) na fala culta carioca, de Pereira (1997) na fala pessoense e de Marques (2006) entre paraibanos no Rio de Janeiro.<sup>5</sup> Marques (2006) assinala que as proporções

---

<sup>5</sup>Pereira (1997) se refere ao trabalho de mestrado dessa autora, ao qual não se teve acesso nesta pesquisa; as proporções de cada variante reportadas por Marques (2006), no entanto, correspondem exatamente aos dados reportados em Pereira (2010), sua tese de doutorado. Note-se que as proporções das vogais média-baixas da fala carioca aí reportadas (4%) diferem daquela indicada neste relatório (2,8%), pois Marques (2006) considera a distribuição de dados antes de uma série de exclusões de contextos categóricos descritas em Yacovenco (1993).



das variantes média-baixas [ɛ] e [ɔ] são maiores em João Pessoa do que na fala de migrantes (Cf. para [ɛ], 44% vs. 39%; para [ɔ], 42% vs. 33%), o que seria indício de sua acomodação aos padrões da comunidade anfitriã; para a variável /e/, há também diminuição dos índices de alçamento (Cf. 34% vs. 20%). O emprego dessas variantes diminui em favor das realizações [e] e [o], as quais, como visto em Yacovenço (1993), são as formas mais frequentes no Rio de Janeiro.

**Tabela 1:** Comparação entre proporções de variantes de vogais médias pretônicas em três estudos (adaptada de Marques 2006 : 86)

	/e/			/o/		
	ɛ	e	i	ɔ	o	u
Yacovenço (1993)	4%	64%	32%	4%	66%	30%
Pereira (1997)	44%	21%	34%	42%	22%	35%
Marques (2006)	39%	41%	20%	33%	31%	36%

Na análise binária de 765 dados de [e] vs. [ɛ], Marques (2006) constatou a influência das variáveis Vogal da Sílabla Seguinte, Tonicidade da Sílabla Seguinte, Tempo de Residência no Rio de Janeiro, Contextos Fonológicos Precedente e Seguinte, Anos de Escolarização e Faixa Etária. Já para a análise de [o] vs. [ɔ], em um total de 418 dados, a autora verificou correlação com as variáveis Vogal da Sílabla Seguinte, Contexto Fônico Seguinte, Tempo de Residência, Faixa Etária, Anos de Escolarização e Contexto Fônico Precedente. Assim como em Pereira (2010), a Vogal da Sílabla Seguinte se mostra preponderante para os padrões de abaixamento. Quanto às variáveis sociais, favorecem a realização [-alta] de /e/ os migrantes com menor tempo de residência no Rio de Janeiro (até 4 anos), os menos

escolarizados (de 0 até 8 anos de escolarização) e os mais velhos; de modo semelhante, favorecem o abaixamento de /o/ os migrantes recentes e os mais velhos.<sup>6</sup>

Em todos esses estudos, a variável dependente *altura da vogal pretônica* foi codificada nominalmente, em categorias discretas (vogal alta, média-alta e média-baixa), e analisada em modelos de regressão logística em versões diversas do programa VARBRUL. No presente estudo, diferentemente, a variável dependente é analisada de acordo com as medidas de F1 em modelos de regressão linear na plataforma R, de modo que os resultados obtidos são comparáveis apenas de modo global: interessa verificar se as mesmas variáveis independentes se mostram correlacionadas e qual a direção da correlação. Outra diferença do presente estudo é a consideração de variáveis aleatórias, como o Informante e o Item Lexical, que, como se verá na seção 2.4, têm papel fundamental para entender padrões gerais de variação: muito da variabilidade nos dados não se deve a amplas categorias como “sexo/gênero” ou “distância da sílaba tônica”, mas ao comportamento dos indivíduos e de itens lexicais específicos. A próxima seção descreve a organização da amostra e os métodos empregados neste estudo.

## 2.2 Organização e caracterização das amostras

O projeto propôs a análise de vogais médias pretônicas na fala de seis migrantes paraibanos residentes no Rio de Janeiro e seis paraibanos residentes em São Paulo, em comparação com padrões de fala de um mesmo número de

---

<sup>6</sup>No entanto, aqui se questiona os resultados para a variável Faixa Etária, uma vez que os dados de Marques (2006) indicam clara interação entre essa variável e as demais variáveis sociais.

indivíduos que fossem nativos e tivessem perfis sociais semelhantes (sobretudo quanto ao grau de escolarização e faixa etária).

Os conjuntos de gravações com falantes nativos foram organizados a partir das amostras já disponíveis do PEUL, coletadas no início dos anos 2000 (Paiva & Duarte, 2003),<sup>7</sup> e do Projeto SP2010, cuja coleta ocorreu no biênio 2012–2013 (Mendes & Oushiro, 2012).<sup>8</sup> As gravações selecionadas se encontram no Quadro 2.<sup>9</sup>

	<b>Informante</b>	<b>Sexo</b>	<b>Idade</b>	<b>Escolaridade</b>
PEUL	16-Car	M	48	Fundamental II
	17-Sim	F	27	Fundamental II
	18-Luc	F	49	Fundamental II
	19-Jor	M	37	Médio
	20-Rei	M	47	Fundamental II
	22-Ana	F	34	Médio
SP2010	AliceM	F	30	Médio
	AnaS	F	32	Médio
	RenataC	F	35	Médio
	RobsonF	M	31	Médio
	NelsonF	M	37	Médio
	MauricioB	M	36	Médio
	LucianoT	M	39	Médio

**Quadro 2:** Informantes nativos das amostras PEUL e SP2010

De modo semelhante, também já se dispunha de gravações com migrantes paraibanos na cidade de São Paulo, coletadas em 2009 pelo Grupo de Estudos

<sup>7</sup>Disponível em <http://www.letras.ufrj.br/peul/censo%202000.html>. Último acesso em 14/04/2016. Embora essa amostra se distancie cronologicamente das demais entre 9 até 15 anos, o fato de o estudo de Yacovenco (1993) ter apontado para um caso de variação estável permite supor que os padrões de variação no Rio de Janeiro não mudaram substancialmente desde então.

<sup>8</sup>Disponível em <http://projetosp2010.fflch.usp.br/corpus>. Último acesso em 14/04/2016.

<sup>9</sup>Nas gravações do PEUL, os informantes são identificados pelo número da gravação e as iniciais de seus nomes; na amostra SP2010, os falantes são identificados por pseudônimos.

e Pesquisa em Sociolinguística da USP (Amostra PBSP). Coube a este projeto, portanto, a coleta de gravações com paraibanos residentes no Rio de Janeiro. Esses informantes foram contactados por intermédio de Marcelo Melo, doutorando do Programa de Pós-Graduação em Linguística da UFRJ, dentre os alunos do curso de Educação de Jovens e Adultos da Escola Politécnica de Saúde Joaquim Venâncio – Fiocruz, localizada no bairro de Manguinhos. As gravações foram realizadas no auditório da Fiocruz, com gravador TASCAM DR-100 e microfones Sennheiser, ao longo dos meses de junho e julho de 2015.

O intuito inicial de se gravarem seis informantes foi parcialmente atendido; devido a um problema técnico na gravação de um casal de falantes, sua fala não foi registrada e tal entrevista teve que ser descartada. A Amostra PBRJ conta, ao final, com cinco gravações. Para compensar a perda desse registro, incluiu-se um informante adicional na Amostra PBSP. O Quadro 3 resume as informações sobre os participantes paraibanos: seu pseudônimo, sexo/gênero, idade, grau de escolaridade, idade no momento da migração e há quantos anos vivem nas respectivas capitais do sudeste.<sup>10</sup>

A amostra completa, assim, constitui-se de 25 informantes (13 nativos e 12 paraibanos), balanceados quanto ao sexo/gênero (12 mulheres e 13 homens) e com perfis sociais semelhantes: idade de 30 a 49 anos e escolaridade até, no máximo, o Ensino Médio. A padronização quanto ao perfil social se deve ao número relativamente pequeno de informantes cuja fala poderia ser coletada e analisada dentro do período previsto de desenvolvimento deste projeto (um ano); uma amostra com perfis mais variados não permitiria chegar a

---

<sup>10</sup>Ver também o perfil estendido dos informantes na página 84.

	<b>Informante</b>	<b>Sexo</b>	<b>Idade</b>	<b>Escolaridade</b>	<b>Migração</b>	<b>Anos no SE</b>
PBRJ	RosalindaF	F	43	Fundamental I	25	18
	JoanaA	F	41	Fundamental I	12	29
	TatianeE	F	29	analfabeta	19	10
	RonaldoS	M	31	Fundamental II	16	15
	JoãoC	M	42	Fundamental I	17	25
PBSP	MartaS	F	32	Fundamental I	18	14
	JosaneV	F	37	Médio	21	16
	MarinalvaS	F	39	Fundamental I	18	21
	JosuéO	M	30	Médio	15	15
	JoãoS	M	31	Fundamental I	14	17
	HenriqueA	M	42	Médio	17	25
	PedroC	M	36	Fundamental I	13	23

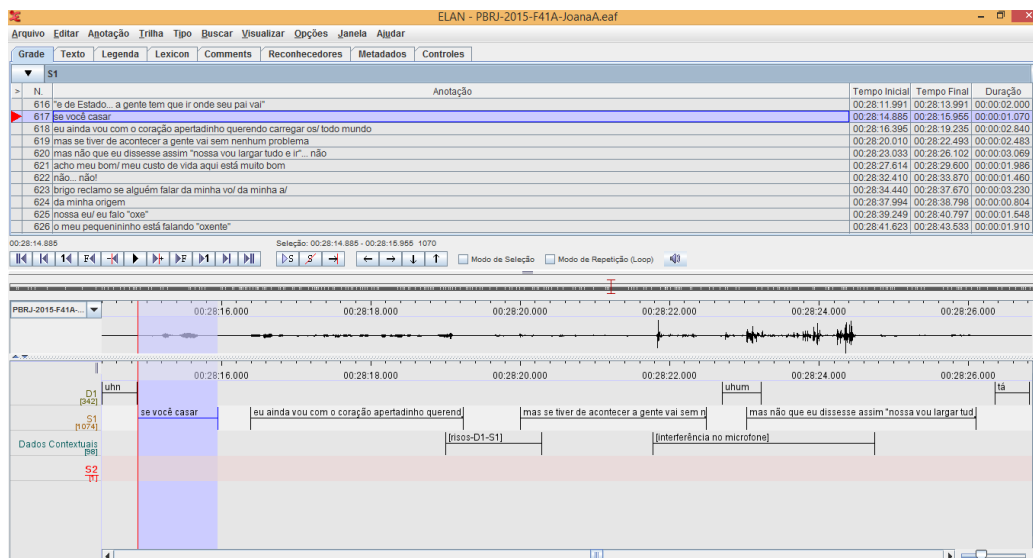
**Quadro 3:** Informantes paraibanos das amostras PBRJ e PBSP

conclusões minimamente confiáveis quanto ao papel de diferentes variáveis sociais em processos de acomodação dialetal – é difícil, por exemplo, inferir o papel de *idade de migração* e *anos de residência no sudeste* caso se compare a fala de um migrante de 60 anos que se deslocou há mais de 40 com a fala de um migrante de 21 anos que, forçosamente, tenha migrado há apenas poucos anos. Trata-se, portanto, de amostras relativamente homogêneas de falantes adultos com grau baixo ou médio de escolaridade (e, inferencialmente, de classe social média ou média baixa).

Embora se intencionasse inicialmente coletar entrevistas apenas com falantes que houvessem migrado com mais de 18 anos – ou seja, na fase adulta –, o contato com a comunidade de paraibanos no Rio de Janeiro e a inspeção de gravações coletadas em São Paulo revelaram uma realidade distinta: muitos dos informantes em potencial haviam se deslocado antes dos 18, tão cedo quanto aos 12 ou 13 anos, em busca de melhores condições de vida. Diante disso, decidiu-se adotar como critério o fato de haverem migrado

por conta própria, e não em decorrência de uma decisão dos pais ou de outros familiares (ainda que pudessem tê-los nas cidades de destino).

Todas as entrevistas foram transcritas ou ajustadas ao programa ELAN (Hellwig & Geerts, 2013), que permite (i) o alinhamento entre transcrição e gravação, com marcação do tempo inicial e final de cada trecho de fala; (ii) a separação dos enunciados de diferentes participantes, além da possibilidade de realizar anotações diversas, como dados contextuais, em trilhas distintas; e (iii) a exportação para o formato .TextGrid do Praat (ver Figura 1). Tais recursos facilitam a análise de variáveis sociolinguísticas, sobretudo de natureza fonética (Oushiro, 2014). As gravações da Amostra SP2010 já se encontravam nesse formato, e aquelas do PEUL já dispunham de transcrição no programa similar EXMARaLDA (Schmidt & Worner, 2014). Durante a execução desse projeto, foram transcritas no ELAN as 12 gravações com paraibanos (PBRJ e PBSP).



**Figura 1:** Transcrição de gravações no programa ELAN

Além disso, obteve-se uma cópia de gravações coletadas pelo Projeto ValPB,<sup>11</sup> que serão analisadas em futuros passos desta pesquisa como parâmetros adicionais aos padrões de fala dos migrantes paraibanos.

## **2.3 Desenvolvimento de novas ferramentas e procedimentos metodológicos para o estudo das vogais pretônicas**

### **2.3.1 Transcritor *silac* e o *script silacpret***

Nos primeiros meses do pós-doutorado, desenvolveu-se um *script* na plataforma R, o transcritor *silac* (Oushiro, 2015), que faz a silabificação, acentuação e transcrição fonológica de textos transcritos em ortografia padrão do Português Brasileiro (PB). Um exemplo do resultado da aplicação do *script* se encontra em (1), que apresenta a transcrição ortográfica à esquerda e a respectiva transcrição fonológica à direita:

---

<sup>11</sup>Disponível em [projetoalpb.com.br/projetoalpb](http://projetoalpb.com.br/projetoalpb). Último acesso em 16/04/2016.

(1)

**Transcrição ortográfica**

ai eu nem peguei o meu  
vou pegar porque se tocar  
oi  
{I.} minha filha  
transferir o quê?  
eu não acredito  
que chatinha hein  
ó então fica com o meu celular  
que acho que até oito horas eu termino  
se tocar você atende pra mim  
não  
pode atender  
ah eu não não não  
(pede pra ver quem é)  
o telefone sem fio está lá?  
está com você? fica com você então

**Transcrição fonológica**

aj 'ew 'nE pe-'gej o 'mew  
'vow pe-'gar 'por-ke se to-'kar  
oj  
i 'mi-Na 'fi-La  
tRAs-fe-'Rir o 'ke  
ew 'nAw a-kRe-'di-to  
ke Sa-'ti-Na e-'I  
o E-'tAw 'fi-ka kO o 'mew se-lu-'lar  
ke 'a-So ke a-'te 'oj-to 'o-Ras 'ew ter-'mi-no  
se to-'kar vo-'se a-'tE-de pRa 'mI  
nAw  
'po-de a-tE-'der  
a 'ew 'nAw 'nAw 'nAw  
'pe-de pRa 'ver 'kE 'e  
o te-le-'fo-ne 'sE 'fio es-'ta 'la  
es-'ta kO vo-'se 'fi-ka kO vo-'se E-'tAw

De modo geral, as regras de transcrição fonológica seguem as convenções do IPA (International Phonetic Alphabet). Entretanto, foram estabelecidas certas normas específicas, sobretudo quanto a caracteres especiais:

- (i) vogais nasais são representadas por letras maiúsculas: a-'sI para *assim*, E-tAw para *então*
- (ii) diferenciam-se o “r” forte (como em *arroz*, *rato*, representado por “h”:



- a-'hos, 'ha-to), o “r” fraco (como em *caro*, *prato*, representado por “R”: 'ka-Ro, 'pRa-to) e o “r” em coda silábica (como em *quarteirão*, representado por “r”: kwar-tej-'RAw)
- (iii) os dígrafos “nh” e “lh” são representados por “N” e “L” respectivamente:  
'mi-Na para *minha*, 'fa-La para *falha*
- (iv) o fonema /ʒ/ é representado por “Z”: 'ZE-te para *gente*, 'Zi-Ra para *gira*
- (v) o fonema /ʃ/ é representado por “S”: 'Sow para *show*, paj-'SAw para *paixão*

Na versão atual, o transcritor ainda sofre algumas limitações: (i) palavras estrangeiras normalmente não são transcritas corretamente, uma vez que as convenções ortográficas de outras línguas mormente não seguem aquelas do PB; e (ii) o transcritor não discrimina a diferença entre vogais média-baixas /ɛ, ɔ/ e média-altas /e, o/. Para certos itens lexicais, tal diferenciação pode ser feita apenas se se tem informação de sua classificação morfológica e função sintática numa dada sentença – p.ex., “almoço” como substantivo ([mo]) ou como verbo ([mɔ]). Ainda que outros tantos casos sejam previsíveis a partir da ortografia (p.ex., para *pastel* /pas-'tɛw/), preferiu-se padronizar a transcrição de vogais médias apenas como “e” e “o” até que se encontre uma solução para todos os casos.

Cabe ressaltar que, em se tratando de uma transcrição fonológica, o transcritor *não* discrimina fenômenos variáveis do PB, como o alçamento de vogais postônicas (p.ex. /o/ → [ʊ], como em *rato*, *cavalo*), a realização variável do segmento /r/ em coda silábica (p.ex. [mar] para *mar*), a palatalização de /t/ e /d/ quando antecedem uma vogal alta anterior (p.ex. [tʃia] para



lexical do qual foi extraída uma ocorrência ou a extração de medidas de formantes (F1 e F2).

Uma dessas ferramentas é o *plugin* EasyAlign (Goldman, 2011), que se instala no programa Praat. O EasyAlign realiza, em princípio, três tarefas: (i) a macrossegmentação de sequências de transcrição, i.e., identifica, em uma cadeia sonora, o ponto inicial e final de sentenças transcritas; (ii) a fonetização, i.e., a transcrição fonológica a partir de uma transcrição ortográfica; e (iii) a segmentação fônica, identificando o tempo inicial e final de cada segmento junto à onda sonora. A aplicação da função (i) era desnecessária, uma vez que as transcrições já haviam sido feitas no ou ajustadas ao programa ELAN (ver Figura 1 acima).<sup>12</sup> A função (ii) realiza tarefa similar ao *script silacpret* acima descrito mas, como não realiza a marcação de vogais pretônicas, deu-se preferência à utilização do *script* de autoria própria. Desse modo, do EasyAlign, foi utilizada somente a função (iii), a segmentação fônica. A partir de tal segmentação, o EasyAlign também identifica o ponto inicial e final de cada item lexical dentro da cadeia sonora.

A segmentação fônica e lexical, no entanto, tampouco é sem problemas. Intenciona-se realizar, futuramente, uma avaliação mais sistemática dos tipos de “erros” de segmentação feitos pelo *plugin* mas, por ora, é possível afirmar que o EasyAlign tende a cometê-los em trechos com muitos ruídos de fundo, em que há sobreposição de vozes de interlocutores, em falas rápidas e quando

---

<sup>12</sup>No entanto, para fins de avaliação do *plugin*, também se realizaram alguns testes assistemáticos, partindo-se de transcrições sem marcações de tempo a fim de verificar o quanto o EasyAlign “acertava” na tarefa de macrossegmentação de sentenças; as taxas de erro foram tão grandes que logo se abandonou a perspectiva de utilizá-lo em estudos futuros para o alinhamento automático de transcrições com gravações. É possível que sua taxa de acerto seja bastante maior em gravações mais curtas, de um único falante e em situações controladas de fala, como a leitura de um texto.

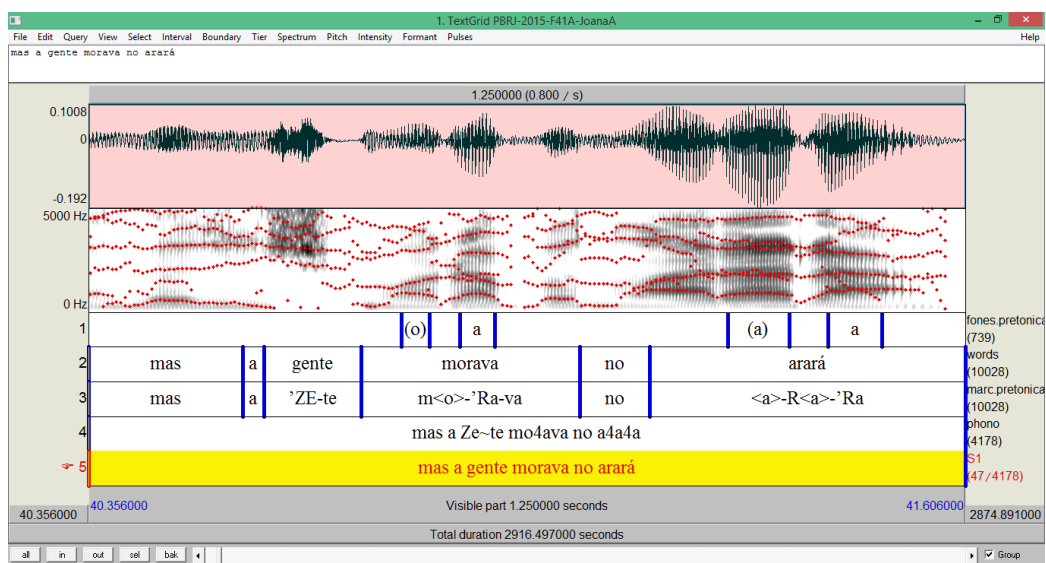
há produções orais como risos e tosses. Contudo, como a segmentação lexical já permitia a localização bastante próxima, ainda que por vezes imprecisa, de palavras que continham vogais pretônicas, a aplicação da função (iii) do EasyAlign foi bastante útil para agilizar a localização e a marcação dos segmentos-alvo deste estudo.

A Figura 2 mostra um arquivo de som junto à sua transcrição no programa Praat após a codificação de vogais. A trilha 5, “S1”, corresponde à fala do informante tal como havia sido exportada do programa ELAN.<sup>13</sup> A trilha 4, “phono”, corresponde à transcrição fonológica do EasyAlign; note-se que as convenções empregadas pelo *plugin* diferem daquelas adotadas para o *silacpret* (por exemplo, o “r” intervocálico é indicado por “4” em vez de “R”). As trilhas 2 e 3 correspondem à segmentação lexical, na forma ortográfica e fonológica, respectivamente. A trilha 3 foi criada por meio da aplicação do *silacpret*, no R, a fim de identificar os segmentos de vogais pretônicas (notem-se os segmentos indicados por “< >” nessa trilha). A partir da trilha 3, realizaram-se buscas por “<e>” e “<o>” para localizar rapidamente as ocorrências de vogais médias pretônicas; tais vogais foram então segmentadas manualmente na trilha 1. Além das vogais médias, foram marcadas as demais vogais pretônicas [i, a, u], a fim de que se pudesse compor todo o quadro vocálico pretônico de cada falante, e as vogais tônicas de cada item lexical (indicadas sem os parênteses, para diferenciá-las das pretônicas), para que se pudesse analisá-las como variável independente.<sup>14</sup>

---

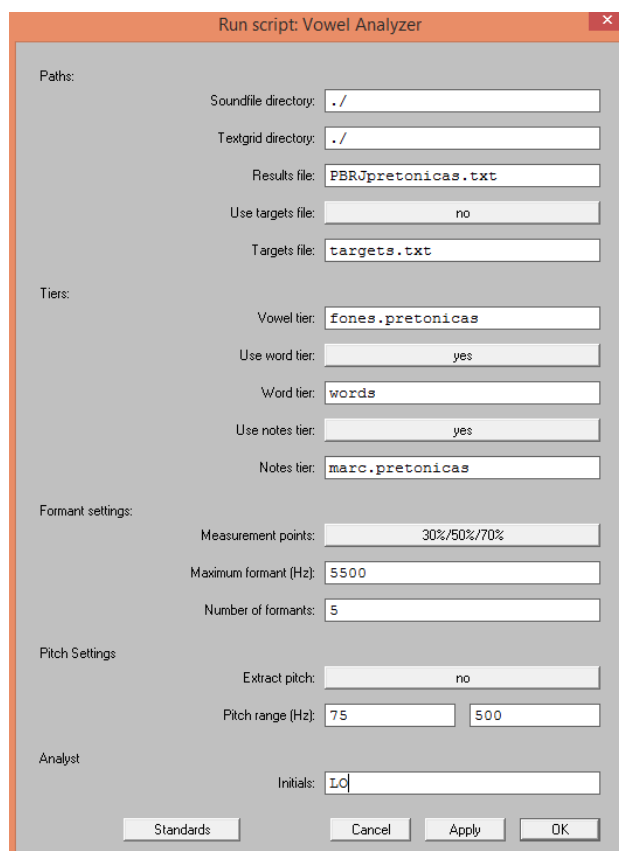
<sup>13</sup>As trilhas com a fala do documentador e de dados contextuais foram descartadas nesse passo da análise.

<sup>14</sup>Os critérios para a inclusão de dados são descritos mais detalhadamente na seção 2.4. Não se incluíram, por exemplo, ocorrências de vogais pretônicas em sílabas sem ataque silábico, como o primeiro /a/ de “Arará”, à semelhança do que fez Pereira (2010).



**Figura 2:** Transcrição no Praat após a aplicação do EasyAlign e marcação de vogais pretônicas e tônicas na trilha 1

Após a segmentação de vogais pretônicas e tônicas em todas as gravações, aplicou-se o *script Vowel Analyzer* (Riebold, 2013) para extração automática de medições de F1, F2, duração de segmentos e respectivos itens lexicais. A Figura 3 mostra a janela que se abre ao rodar esse *script*; nela, especificam-se: (i) os diretórios em que se encontram os arquivos de gravação e de transcrição no formato .TextGrid (a notação “./” indica que é o mesmo diretório em que está o *script*); (ii) o nome do arquivo de resultados (neste caso, “PBRJpretonicas.txt”); (iii) se se deve utilizar um arquivo de palavras-alvo (neste caso, não); (iv) a “Vowel tier”, trilha que contém a marcação de vogais; (v) a “Word tier”, trilha que contém os itens lexicais; (vi) se se deve extrair um “Notes tier” (neste caso, extraíram-se as informações da trilha de transcrição fonológica); (vii) os pontos de medição de formantes (escolheu-se a opção “30%/50%/70%”); e mantiveram-se as demais opções no formato *default*.



**Figura 3:** Janela do *script* VowelAnalyzer

O arquivo gerado pelo *script* Vowel Analyzer é uma planilha de dados que contém o nome do arquivo de que se extraiu cada dado (o que também identifica o falante); o item lexical (extraído da trilha “words”); a transcrição fonológica (extraída da trilha “marc.pretonicas”); a vogal (pretônicas e tônicas, extraídas da trilha “fones.pretonicas”); o tempo inicial, final e a duração de cada vogal; e as medidas de F1 e de F2, extraídas em 30%, 50% e 70% dentro do intervalo de cada vogal (ver Figura 4).

The image shows a spreadsheet with columns labeled A through J. The data includes phonetic transcriptions, vowel symbols, and numerical values representing time and duration. The first row is the header for 'Informante', and subsequent rows list various phonetic segments and their corresponding acoustic parameters.

A	B	C	D	E	F	G	H	I	J	
Informante	Word	Transc Fon	Vowel	Begin Time s	End Time s	Duration ms	F1 30	F1 50	F1 70	
2	PBRJ-2015-F29A-TatianeE	isapé	s<a>-pe	(a)	4.97366304066983	5.0609176305826935	87.25458991286317	652.5296951478476	714.6533499494669	736.1914022904334
3	PBRJ-2015-F29A-TatianeE	isapé	s<a>-pe		5.233304417948338	5.297957583846134	64.65316589779668	668.4281456433088	669.6271821704013	675.8347022904334
4	PBRJ-2015-F29A-TatianeE	trabalhava	tR<a>-b<a>-La-va	(a)	8.91212581906962	9.957207357191267	45.0815381216465	662.8191621021526	683.832737538856	681.993522274189
5	PBRJ-2015-F29A-TatianeE	trabalhava	tR<a>-b<a>-La-va	(a)	9.026283907538952	9.064094229834526	37.81032295573684	712.9134664765102	725.2277130021671	735.184869535877
6	PBRJ-2015-F29A-TatianeE	trabalhava	tR<a>-b<a>-La-va	a	9.110994939878349	9.135767136885916	24.779643007566903	783.9785784147651	778.9712162346739	773.3721817288991
7	PBRJ-2015-F29A-TatianeE	morar	m<o>-Rar	(o)	40.497711840720754	40.52194922680767	24.2373860869165	633.5106090717868	630.6128095573447	640.503365577577
8	PBRJ-2015-F29A-TatianeE	morar	m<o>-Rar	a	40.613044980880034	40.648694857403115	35.64987652308105	739.804815736762	795.8449708568387	952.3526260971366
9	PBRJ-2015-F29A-TatianeE	morar	m<o>-Rar	(o)	49.16242891436035	49.187389797192594	24.92988283224662	530.1398112979653	566.8113240331385	586.4570258785349
10	PBRJ-2015-F29A-TatianeE	morar	m<o>-Rar	a	49.25212618949411	49.28354641965005	31.420230155937645	726.143436232674	733.2220738336305	728.6498622231213
11	PBRJ-2015-F29A-TatianeE	paraitano	p<a>-R<a>-ba-no	(a)	57.70730367255439	57.73701178229496	29.708110375096908	657.5874494184791	699.987750849075	757.38368844220
12	PBRJ-2015-F29A-TatianeE	paraitano	p<a>-R<a>-ba-no	a	57.80059853185019	57.83383146758205	33.23283574185726	731.6076700049715	727.5406605469545	703.986902346411
13	PBRJ-2015-F29A-TatianeE	laqoa	l<a>-go-a	(a)	60.369642923252438	60.442569139808704	72.91990728432296	662.6843193049879	689.3960892389617	684.7537815462989
14	PBRJ-2015-F29A-TatianeE	laqoa	l<a>-go-a	o	60.53832692265708	60.5633032710976	44.71340445267913	461.00370648900616	452.8660048373696	425.61161584153
15	PBRJ-2015-F29A-TatianeE	trabalhava	tR<a>-b<a>-La-va	(a)	73.29956546453325	73.36337038340703	63.804918873785255	811.2371333575526	774.394740212872	752.958607107132
16	PBRJ-2015-F29A-TatianeE	trabalhava	tR<a>-b<a>-La-va	(a)	73.44517156145035	73.4909802211546	45.80865870424952	800.1240805815087	791.8960583976397	804.3824136121787
17	PBRJ-2015-F29A-TatianeE	trabalhava	tR<a>-b<a>-La-va	a	73.55432093975965	73.58876234588912	34.44140513247271	763.104421527039	761.608821234310	757.5080380580015
18	PBRJ-2015-F29A-TatianeE	trabalhando	tR<a>-b<a>-LA-do	(a)	75.6694082918372	75.73321321071097	63.804918873771044	686.4731294538769	713.701919694658	703.865884444483
19	PBRJ-2015-F29A-TatianeE	trabalhando	tR<a>-b<a>-LA-do	a	75.84259201569422	75.87824189221732	35.64987652309526	719.2805209292704	731.7884962658058	734.1640822983335
20	PBRJ-2015-F29A-TatianeE	morar	m<o>-Rar	(o)	83.40669186501059	83.299078183981	23.2195953852047	632.7901112118628	637.3994044484841	626.376043695785
21	PBRJ-2015-F29A-TatianeE	morar	m<o>-Rar	a	83.47619664917218	83.52816087596854	51.96422679635759	668.6584348167081	660.7605034840034	646.4561998432595
22	PBRJ-2015-F29A-TatianeE	fiquei	f<i>-kej	(i)	94.07170912265855	94.08763654780137	15.927425142820084	374.9703110950117	382.40957359289007	384.33032362123
23	PBRJ-2015-F29A-TatianeE	fiquei	f<i>-kej	ej	94.22457951882194	94.28049030352558	60.4235195306444	484.1361194856506	505.01994444960644	526.994725146500
24	PBRJ-2015-F29A-TatianeE	batia	b<a>-tia	(a)	112.18606927871995	112.26522443320239	80.16515448244377	752.9650123323308	794.5216569012333	778.5720957910086
25	PBRJ-2015-F29A-TatianeE	batia	b<a>-tia	i	114.1751361524222	114.2436933303039	68.55717778867643	408.9569132916495	415.1767224223586	418.70868445439726
26	PBRJ-2015-F29A-TatianeE	vivia	v<i>-via	(i)	114.40930682450141	114.42924658594653	19.939761445115778	410.8359173934763	413.6181498019865	417.8469771496176
27	PBRJ-2015-F29A-TatianeE	vivia	v<i>-via	i	114.40930682450141	114.42924658594653	19.939761445115778	410.8359173934763	413.6181498019865	417.8469771496176
28	PBRJ-2015-F29A-TatianeE	morar	m<o>-Rar	(o)	152.26819644365227	152.2889713460125	20.774902360216174	649.0653411232862	657.1056050441438	657.8158943982302
29	PBRJ-2015-F29A-TatianeE	morar	m<o>-Rar	a	152.35098234698756	152.3894916429186	38.06581730429773	582.5370892874355	600.7864327884845	672.8739491850772
30	PBRJ-2015-F29A-TatianeE	catorze	k<a>-tor-ze	(a)	155.2003107550861	155.25023946456827	49.92870948217387	723.776813318438	730.892659539296	897.1893088230444
31	PBRJ-2015-F29A-TatianeE	catorze	k<a>-tor-ze	o	155.38951731166895	155.4197290714343	30.21175976533641	539.125252123327	522.599111424698	508.1530603947224

Figura 4: Arquivo de dados gerado pelo *script* Vowel Analyzer para o Praat

### 2.3.3 Codificação automática de variáveis

Em cima do arquivo gerado pelo Vowel Analyzer, aplicou-se um novo *script* elaborado pela bolsista na plataforma R para a codificação automática das seguintes variáveis independentes:

- Contexto Fônico Precedente
- Contexto Fônico Seguinte
- Vogal da Sílabla Seguinte (/a/, /e/ etc.)
  - (a) F1 da vogal da sílabla seguinte (em Hz)
  - (b) F2 da vogal da sílabla seguinte (em Hz)
- Vogal Tônica (/a/, /e/, /i/ etc. )
- Distância da Vogal Tônica (em número de sílabas)
- Estrutura da Sílabla Pretônica (CV, CVC, CCV etc.)
- Amostra (PBRJ, PBSP, PEUL, SP2010)
- Informante

- Sexo (F, M)
- Idade (em anos)
- Idade de Chegada ao Rio de Janeiro ou São Paulo (em anos)
- Anos de Residência no Sudeste

São oito variáveis linguísticas frequentemente analisadas em estudos sobre vogais médias pretônicas, conforme descrito na seção 2.1, e seis variáveis sociais de interesse a este estudo. As variáveis Contexto Fônico Precedente, Contexto Fônico Seguinte e Estrutura da Sílabas Pretônica foram codificadas a partir dos dados das colunas “Vowel” e “Transc.Fon”: uma vez identificada a vogal pretônica dentro da transcrição fonológica, a identificação do segmento que ocorre logo antes e logo depois é previsível, bem como a estrutura da sílaba que contém a vogal (pela delimitação silábica por meio dos hífen). A codificação automática da Vogal da Sílabas Seguinte foi possível devido ao fato de que se anotaram sistematicamente, na trilha de vogais no Praat, todas as vogais das sílabas seguintes aos segmentos-alvo; tanto a sua identificação (/a/, /i/ etc.) quanto as respectivas medidas de F1 e F2 se referem aos dados da linha seguinte da planilha; por exemplo, para a vogal “a” em “sapé”, na linha 2 da Figura 4, os dados da vogal da sílaba seguinte “é” se encontram na linha 3; para a primeira vogal pretônica “a” em “trabalhava” na linha 4, os dados da vogal da sílaba seguinte “a” se encontram na linha 5. A identificação da Vogal Tônica de cada item lexical é possível por sua marcação na transcrição fonológica pelo símbolo ('); uma vez identificada a sílaba tônica, dela se apagam as consoantes (p.ex., de [so-'fRE-do], identifica-se a tônica [fRE] pela marca de tonicidade e a delimitação silábica pelos hífen, e dela se apagam as consoantes “f” e “R”). Por fim, a distância da vogal pretônica



em relação à vogal tônica é possível por meio da contagem de hífen (p.ex., em “m<o>-'Rar”, como há apenas um hífen entre a vogal pretônica < > e a marcação de tonicidade ', a distância da sílaba tônica é “1”).

Uma vez extraídas e codificadas as informações referentes à vogal da sílaba seguinte e da sílaba tônica, as ocorrências correspondentes a vogais tônicas (identificadas pela ausência de parênteses na coluna “Vowel”) foram apagadas do arquivo de dados, já que não se referem a segmentos-alvo deste estudo.

A codificação de variáveis sociais, por sua vez, é bastante previsível, uma vez que as características de amostra, sexo, idade, idade de chegada etc. são constantes para cada falante.

Como variável dependente deste estudo, foram empregadas as medidas de F1 e F2 de cada vogal, tanto em seu formato bruto, extraído pelo *script Vowel Analyzer*, quanto em formato normalizado. Para tanto, empregou-se a normalização de Lobanov (Lobanov, 1971), por meio da função `norm.lobanov` do pacote `vowels` (Kendall & Thomas, 2015) na plataforma R. A normalização de vogais permite maior comparabilidade dos espaços vocálicos entre diferentes indivíduos ao neutralizar diferenças de medidas de formantes decorrentes de características fisiológicas e de trato vocal (p.ex., mulheres tendem a apresentar medidas mais altas de formantes em comparação com as medidas dos homens para as mesmas vogais). Nas análises da seção 2.4, serão apresentados resultados tanto com vogais normalizadas quanto não normalizadas.

O desenvolvimento de *scripts* como o `silacpret` e o de codificação automática de variáveis, assim como a testagem da aplicação de ferramentas já existentes como o `EasyAlign` e o `Vowel Analyzer`, não constituem tarefas

triviais, mas o esforço sistemático de se buscarem métodos mais eficientes de análises de dados. Nesta pesquisa, as tarefas descritas nesta seção 2.3 ocuparam cerca de seis meses (de junho a novembro/2015), concomitantemente à coleta da amostra PBRJ, transcrição das amostras PBRJ e PBSP, e à análise quantitativa de dados. A dedicação deste tempo à elaboração de procedimentos metodológicos, no entanto, certamente é recompensada pela perspectiva de sua aplicação em estudos futuros e na expansão da atual pesquisa. O transcritor fonológico *silac* poderá ser empregado em diversos estudos sobre variáveis fonéticas, não apenas sobre vogais pretônicas. Os procedimentos e *scripts* para o tratamento de vogais podem ser aplicados tanto na expansão das amostras a serem analisadas por esta pesquisadora, quanto por outros (de fato, os *scripts silac* e *Vowel Analyzer* já vêm sendo empregados por dois alunos do Departamento de Linguística da USP). Adicionalmente, pode-se tentar imaginar quanto tempo se levaria para que fossem codificadas manualmente todas as variáveis independentes e as medidas de F1 e F2 de milhares de ocorrências de vogais pretônicas.<sup>15</sup>

A próxima seção discute os resultados preliminares da análise sobre esses dados, e a seção seguinte vislumbra os passos adicionais a serem tomados nesta pesquisa.

---

<sup>15</sup>Uma descrição passo-a-passo dos procedimentos metodológicos se encontra no Anexo A, p. 82.

## 2.4 Análises dos dados

### 2.4.1 Métodos

Conforme se descreveu acima, a segmentação de vogais pretônicas dentro do programa Praat foi feita manualmente, com auxílio da trilha “marc.pretonicas”, em que as vogais-alvo estavam delimitadas pelos símbolos “< >”. Para uma primeira análise dos dados, decidiu-se examinar a realização de vogais médias pretônicas quanto a sua possível realização com o traço [-alto], típico da região Nordeste. A esta pesquisa, interessa verificar se os migrantes paraibanos, em contato com as variedades carioca e paulistana, tenderiam a realizá-la de modo relativamente mais alto (i.e., com um valor de F1 mais baixo); mais especificamente, interessa identificar quais falantes mais se aproximam dos padrões de uso dos nativos e, uma vez identificados, levantar hipóteses sobre quais fatores podem ter contribuído para maior acomodação dialetal.

Sabendo-se que certas vogais médias pretônicas também podem ser realizadas com o traço [+alto] na variedade paraibana (p.ex. [mi.'ni.nʊ]), o envelope de variação desta análise se delimitou pelos contextos em que há maior tendência à realização [-alta] dessas vogais, de acordo com Pereira (2010): vogais em sílabas pretônicas cuja sílaba seguinte contém uma vogal oral [-alta] /a, ε, ɔ/ (como em *relógio*, *moradia*), um ditongo oral /aw, εw, ɔw/ (como em *geral*, *totalmente*), uma vogal nasal ou nasalizada [-alta] /ã, ã̃, õ/ (como em *sessenta*, *problema*) ou um ditongo nasal /ãw̃/ (como em *impressão*, *portão*). De modo similar a Pereira (2010), incluíram-se apenas vogais em sílabas com ataque silábico (CV, CVC, CCV, CCVC) – ou seja, não se incluem dados como o primeiro /e/ de “elefante” –, e tampouco se

incluíram vogais em hiatos (p.ex. *realmente*), em prefixos (p.ex. *refazer*), e em nomes próprios de pessoas (p.ex. *Renato*, mas se mantiveram dados como *Pernambuco* e *Nordeste*). Evitou-se também recolhê-las quando seguidas de consoantes nasais (p.ex. *remédio*), contexto que desfavorece o traço [-alto], embora algumas dessas ocorrências tenham sido extraídas a fim de garantir um número mínimo de dados por falantes. Para cada um dos 25 informantes, marcaram-se 50 ocorrências de vogais /e/ e /o/ pretônicas ou o máximo possível de dados que cumpriam os critérios acima (ver Tabela 2; alguns falantes têm um menor número de ocorrências).

Além das vogais médias, também foram marcados e extraídos 30 dados (ou tantos quanto de fato ocorreram na gravação) de cada uma das vogais /a/, /i/ e /u/ em posição pretônica, para cada um dos informantes, com o intuito de estabelecer um quadro de seus respectivos espaços vocálicos nessa posição acentual. Para essas vogais, levou-se em consideração apenas o critério de que houvessem ocorrido em sílaba com ataque silábico e fora de hiato.

A partir desses critérios, analisa-se um total de 4.683 ocorrências de vogais pretônicas, das quais 2.449 são de vogais médias /e/ e /o/. As medidas de F1 e F2 foram utilizadas para calcular a posição média de cada vogal por falante e por amostra (ver figuras adiante). Como variável dependente, utilizou-se a medida de F1 (altura) em seu ponto médio (50%); as medidas em 30% e 70% da vogal foram utilizadas para checagem de eventuais valores discrepantes ou atípicos (p.ex., uma medida de F1 de 300 Hz para a vogal /e/) – tais ocorrências se referem, em sua maioria, a medições ocasionalmente equivocadas por parte do Praat e muitas vezes puderam ser ajustadas com

**Tabela 2:** Total de dados analisados

Amostra	Informante	(a)	(e)	(i)	(o)	(u)	<b>Total</b>
PBRJ	JoanaA	30	49	30	49	30	<b>188</b>
	JoaoC	30	50	30	50	30	<b>190</b>
	RonaldoS	30	50	30	50	30	<b>190</b>
	RosalindaF	30	51	30	51	30	<b>192</b>
	TatianeE	30	50	30	50	30	<b>190</b>
PBSP	HenriqueA	30	46	30	18	14	<b>138</b>
	JoaoS	30	51	30	42	30	<b>183</b>
	JosaneV	30	50	30	43	30	<b>183</b>
	JosueO	30	50	30	50	30	<b>190</b>
	MarinalvaS	30	50	30	50	30	<b>190</b>
	MartaS	30	51	30	50	30	<b>191</b>
	PedroC	30	50	30	49	30	<b>189</b>
PEUL	AnaCristina	30	49	30	49	30	<b>188</b>
	Carlos	30	51	30	50	30	<b>191</b>
	Jorge	30	49	30	50	30	<b>189</b>
	Lucia	30	51	30	50	30	<b>191</b>
	Reinaldo	30	50	30	50	30	<b>190</b>
	Simone	30	50	30	50	30	<b>190</b>
SP2010	AliceM	30	50	30	51	30	<b>191</b>
	AnaS	30	50	30	51	30	<b>191</b>
	LucianoT	30	50	30	50	30	<b>190</b>
	MauricioB	30	50	30	49	30	<b>189</b>
	NelsonF	30	50	30	49	30	<b>189</b>
	RenataC	30	50	30	50	30	<b>190</b>
	RobsonF	30	50	30	50	30	<b>190</b>
<b>Total</b>		<b>750</b>	<b>1248</b>	<b>750</b>	<b>1201</b>	<b>734</b>	<b>4683</b>

base nas medições adicionais.<sup>16</sup>

Como visto em 2.1, diversos estudos sociolinguísticos sobre vogais médias pretônicas em diferentes comunidades brasileiras empregam, em sua quase totalidade, o programa de análises estatísticas Varbrul (em suas diferentes versões), com a codificação nominal (e, mormente, de oitiva) das vogais em “alta”, “média”, “média baixa”. Neste trabalho, as vogais foram analisadas em modelos de efeitos mistos de regressão linear na plataforma R, que abre caminho para duas inovações metodológicas no estudo de vogais: a regressão linear se aplica a variáveis dependentes contínuas, como é o caso desta análise (medidas de F1 em Hz das vogais médias pretônicas), e permite captar gradações, diferentemente de variáveis nominais, categóricas (como “vogal média”, “vogal baixa”). Os efeitos mistos, por sua vez, permitem a inclusão de dois tipos de variáveis nos modelos testados: *efeitos fixos* e *efeitos aleatórios*. Os primeiros são variáveis cujos exemplares da amostra são representativos da população amostrada, como o sexo do informante ou o contexto fônico precedente. Tais variáveis e seus respectivos fatores podem ser replicados em outros estudos. Já os efeitos aleatórios se referem a variáveis específicas da amostra analisada, como os indivíduos que foram aleatoriamente selecionados e os itens lexicais que fortuitamente ocorreram na gravação. A inclusão de efeitos aleatórios no modelo estatístico permite verificar se as correlações observadas se devem de fato aos efeitos fixos ou se possivelmente se devem à

---

<sup>16</sup>Também se criou outra variável dependente, “Distância Euclidiana”, que será analisada futuramente. Tal variável mede a distância entre a posição da vogal pretônica e a da vogal da sílaba seguinte; chama-se “euclidiana” pois se calcula como a medida de hipotenusa de um triângulo retângulo que se forma a partir das diferenças entre medidas de F1 e de F2 das duas vogais. A hipótese, neste caso, é que a distância euclidiana entre a pretônica e a vogal [-alta] da sílaba seguinte será menor para os falantes paraibanos do que para os nativos, e que, quanto maior essa distância, maior o grau de acomodação do falante.

contribuição casual de certos indivíduos ou itens lexicais que eventualmente se comportam de modo distinto. Assim, a constatação de correlação com uma variável independente dentro de um modelo de efeitos mistos permite concluir, de modo mais confiável, que tal variável de fato tem um papel na variação, o que gera modelos mais apropriados.

As análises multivariadas incluíram as variáveis linguísticas (i) Contexto Fônico Precedente (codificada, seguindo Pereira 2010, de acordo com o ponto de articulação dental-alveolar, labial, velar, palatal, sibilante<sup>17</sup> e vibrante posterior); (ii) Contexto Fônico Seguinte (também codificada de acordo com ponto de articulação); (iii) F1 da Vogal da Sílab Seguinte (em Hz);<sup>18</sup> (iv) Vogal Tônica; (v) Distância da Vogal Tônica (de 1 a 5 sílabas);<sup>19</sup> e (vi) Estrutura da Sílab Pretônica (CV, CVC, CCV, CCVC). Nos modelos em que se analisou a fala de paraibanos (amostras PBRJ ou PBSP), incluíram-se as variáveis sociais Sexo/Gênero, Idade de Chegada e Anos na Capital; nos modelos que incluíram as amostras PEUL ou SP2010, manteve-se apenas Sexo/Gênero e excluíram-se, evidentemente, as variáveis que se aplicam apenas a paraibanos. Como se verá adiante, no entanto, apenas algumas dessas variáveis (a saber, F1 da Vogal da Sílab Seguinte, Contexto Fônico Precedente e, em um caso, Sexo/Gênero) se mostram significativamente correlacionadas com a altura das vogais médias pretônicas nessas amostras.

---

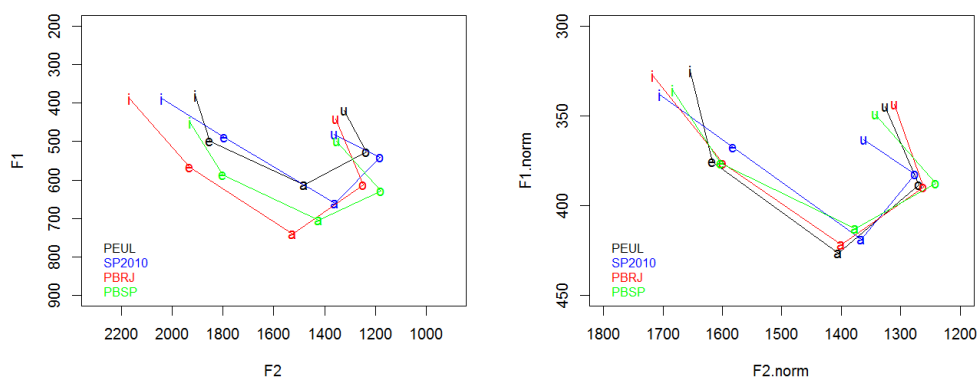
<sup>17</sup>Pode-se questionar a separação de sibilantes /s/ e /z/ das demais consoantes dental-alveolares; para este estudo, no entanto, é preferível manter a comparabilidade dos resultados com os de Pereira (2010).

<sup>18</sup>Note-se que, diferentemente de estudos prévios que se debruçaram sobre “Vogal da Sílab Seguinte” (/a/, /ε/, /ɔ/ etc.), esta pesquisa analisa essa variável de modo contínuo.

<sup>19</sup>Também analisada como variável contínua.

## 2.4.2 Análises globais: comparações entre amostras

A Figura 5 apresenta a posição média das cinco vogais pretônicas em cada uma das quatro amostras – PEUL, SP2010, PBRJ e PBSP –, representadas, respectivamente, pelas cores preta, azul, vermelha e verde. O eixo  $y$  se refere às medidas de F1 e o eixo  $x$  às medidas de F2, e ambos têm seus valores invertidos (valores menores acima e à direita) para refletir adequadamente o posicionamento relativo de cada vogal de acordo com convenção do IPA (p.ex., a vogal [+alta, -anterior] /u/, que tende a valores mais baixos de F1 e de F2, apresenta-se no canto superior direito). À esquerda se encontra o quadro vocálico com vogais não normalizadas (i.e., com suas medidas brutas de F1 e F2), e à direita se apresenta o mesmo quadro em que as medidas dos dois formantes foram normalizadas de acordo com o método de Lobanov (1971).



**Figura 5:** Comparação de médias de F1/F2 de vogais pretônicas nas quatro amostras (esq.: vogais não normalizadas; dir.: vogais normalizadas)

Note-se que a diferença entre a posição média das vogais é muito maior no quadro em que as vogais não foram normalizadas. No quadro à esquerda,



as vogais /e/ e /o/ apresentam-se “mais acima” para as amostras PEUL e SP2010, com valores mais baixos de F1 (próximos a 500 Hz), em comparação com as amostras de paraibanos, em que as medidas de F1 se aproximam de 600 Hz. Isso está de acordo com a expectativa de que os paraibanos produzam vogais médias relativamente mais baixas do que os nativos do sudeste.

Tais diferenças, no entanto, são parcialmente neutralizadas quando se aplica a normalização de Lobanov. Enquanto as vogais /e/ e /o/ dos cariocas parecem se juntar à dos paraibanos, essas vogais permanecem com valores mais baixos de F1 para os paulistanos.

Contudo, cabe examinar se tais diferenças são significativas. Realizaram-se, desse modo, quatro análises iniciais em que se testou se o valor médio de F1 para cada amostra difere para as vogais /e/ e /o/, quando normalizadas ou não.

A Tabela 5 apresenta os resultados do modelo de regressão linear para a vogal pretônica /e/, que incluiu a variável Amostra como efeito fixo e as variáveis Informante e Item Lexical como efeitos aleatórios. Como se trata de modelo ainda pouco usual na Sociolinguística brasileira, cabem inicialmente algumas considerações sobre sua leitura e interpretação. A tabela apresenta, em sua primeira coluna, os níveis da variável independente fixa; o valor de *Intercept* se refere ao nível de referência que, nesta análise, foi estabelecido como “PEUL”. A segunda coluna apresenta as medidas estimadas para a variável dependente (valores de F1), em sua unidade de análise (Hz). Desse modo, a estimativa de valor médio de F1 da vogal /e/ para os cariocas é de 497,2 Hz. Os demais valores de estimativas devem ser lidos *com referência* ao *intercept*; para os paulistanos, a estimativa de valor médio de F1 da vogal

/e/ é 12,4 Hz *abaixo* da estimativa para os cariocas, ou seja,  $497,2 - 12,4 = 484,8$  Hz. A apresentação de resultados na forma de *diferenças* em relação ao *intercept* em vez dos valores estimados em si permite verificar, rapidamente, o quanto tal valor difere de zero: se a estimativa for nula (ou próxima dela), isso significa que as estimativas são praticamente idênticas à do *intercept*; ou, visto de outro modo, que provavelmente não há diferença significativa.

**Tabela 3:** Estimativas (em Hz) de valores de F1 para vogais /e/ pretônicas nas quatro amostras (N = 1.248)

	Estimativa	Erro padrão	Valor de $t$	Significância ( $p$ )	
( <i>Intercept</i> )	497,216	81,157	6,127	<0,001	***
SP2010	-12,382	114,715	-0,108	0,916	
PBRJ	67,896	89,111	0,762	0,465	
PBSP	87,151	86,909	1,003	0,341	

Modelo: `lmer(F1~AMOSTRA+(1|INFORMANTE)+(1|Word), data=dados.e)`

A terceira coluna apresenta o erro padrão, uma medida de dispersão que dá indícios de quão precisa é a estimativa. Quanto maior o valor de erro padrão, maior é a variabilidade nas medições. Na Tabela 3, verifica-se que o erro padrão foi razoavelmente alto na amostra SP2010, o que indica que houve maior variabilidade nas medidas de F1 da vogal /e/ para os participantes paulistanos. O valor de  $t$ , na quarta coluna, é calculado pela razão Estimativa / Erro Padrão (p.ex., para PBRJ:  $67,896 / 89,111 = 0,762$ ). Tal valor é usado para calcular um valor de significância, que mede a probabilidade de se observar tal distribuição em caso de a hipótese nula ser verdadeira; por convenção, estabelece-se um limite de até 5% (0,05) para que um resultado seja considerado significativo. É importante notar que, na coluna de significância, a primeira linha sempre terá um valor de  $p$

bem abaixo de 0,05, pois tal medida apenas avalia a probabilidade de se ter observado tal valor de *intercept* (no exemplo, 497,2 Hz) caso se esperasse um valor zero; ora, não se esperaria que o valor de F1 para a vogal /e/ fosse de zero Hz, de modo que o valor de significância da primeira linha das tabelas não é relevante para esta análise (ela é apresentada apenas por convenção).

A Tabela 3 mostra, assim, que a estimativa de medida de F1 da vogal /e/ para os paulistanos é -12,4 Hz em relação à dos cariocas (i.e., um vogal mais alta), mas que tal diferença não é significativa. Por sua vez, os paraibanos, tanto no Rio de Janeiro quanto em São Paulo, apresentam estimativas de F1 da vogal /e/ bastante acima da dos cariocas: 67,9 Hz a mais para os paraibanos da amostra PBRJ (i.e., 565,1 Hz), e 87,2 Hz a mais para os paraibanos da amostra PBSP (i.e., 584,4 Hz). Essas medidas refletem o posicionamento mais “abaixo” da vogal /e/ para paraibanos na Figura 5. No entanto, o modelo de regressão logística calcula uma probabilidade bem acima do limite de 0,05 de que se tenham observado tais medidas em caso de a hipótese nula ser verdadeira, ou seja, diferenças não significativas. Mesmo que a vogal /e/ de paraibanos se apresente, no quadro à esquerda da Figura 5, abaixo da vogal de cariocas e paulistanos, não se pode afirmar que esses falantes de fato produzam o /e/ pretônico significativamente mais baixo.

De modo bastante semelhante, a Tabela 4, que apresenta os resultados de regressão linear para a vogal /o/, mostra que a estimativa de valor de F1 dessa vogal para os paulistanos é bastante próxima à dos cariocas (diferença de 5,6 Hz) e que tal diferença não é significativa. Adicionalmente, as estimativas de F1 da vogal /o/ para os paraibanos também se apresentam bastante mais altas do que para cariocas e paulistanos (71,6 Hz a mais para paraibanos

no Rio de Janeiro e 86,5 Hz a mais para paraibanos em São Paulo) mas, igualmente, tais diferenças não são avaliadas como significativas pelo modelo.

**Tabela 4:** Estimativas (em Hz) de valores de F1 para vogais /o/ pretônicas nas quatro amostras (N = 1.201)

	Estimativa	Erro padrão	Valor de $t$	Significância ( $p$ )	
<i>(Intercept)</i>	533,131	81,293	6,558	<0,001	***
SP2010	5,625	114,823	0,049	0,962	
PBRJ	71,639	89,19	0,803	0,442	
PBSP	86,534	87,073	0,994	0,345	

Modelo: `lmer(F1~AMOSTRA+(1|INFORMANTE)+(1|Word), data=dados.o)`

As Tabelas 5 e 6 apresentam o resultado de análises similares, em que se analisaram as medidas de F1 normalizadas (correspondentes ao quadro à direita da Figura 5). Nessas duas análises, percebe-se que as estimativas de F1 em Hz são menores, assim como as medidas de erro padrão – com efeito, a técnica de normalização de vogais visa justamente à eliminação de diferenças que possam ser atribuídas aos indivíduos.<sup>20</sup> Caso houvesse diferenças significativas entre amostras na análise de vogais normalizadas, seria possível concluir, mais confiavelmente, que os padrões de paraibanos e nativos do sudeste de fato são divergentes.

Contudo, antes de interpretar que os paraibanos das amostras PBRJ e PBSP se acomodaram aos padrões do sudeste pelo fato de não haver diferenças significativas entre as estimativas, devem ser levados em conta três outros motivos possíveis: os valores relativamente altos de erro padrão, tanto na análise de /e/ quanto de /o/ não normalizados, para todas as

<sup>20</sup>Analiticamente, tal procedimento se assemelha à inclusão da variável Informante no modelo estatístico.

**Tabela 5:** Estimativas (em Hz) de valores de F1 (normalizados) para vogais /e/ pretônicas nas quatro amostras (N = 1.248)

	Estimativa	Erro padrão	Valor de $t$	Significância ( $p$ )	
<i>(Intercept)</i>	375,444	3,304	113,619	<0,001	***
SP2010	-8,735	4,566	-1,913	0,188	
PBRJ	0,530	3,897	0,136	0,900	
PBSP	0,158	3,730	0,042	0,969	

Modelo: `lmer(F1.norm~AMOSTRA+(1|INFORMANTE)+(1|Word), data=dados.e)`

**Tabela 6:** Estimativas (em Hz) de valores de F1 (normalizados) para vogais /o/ pretônicas nas quatro amostras (N = 1.201)

	Estimativa	Erro padrão	Valor de $t$	Significância ( $p$ )	
<i>(Intercept)</i>	389,134	6,941	56,061	<0,001	***
SP2010	-7,361	9,729	-0,757	0,478	
PBRJ	-2,384	7,726	-0,309	0,767	
PBSP	-3,573	7,552	-0,473	0,651	

Modelo: `lmer(F1.norm~AMOSTRA+(1|INFORMANTE)+(1|Word), data=dados.o)`

amostras, o que indica ampla variabilidade entre indivíduos ou itens lexicais; a inclusão das variáveis aleatórias Informante e Item Lexical no modelo – o que, novamente, aponta para a importância dessas variáveis, mais do que o papel do efeito fixo Amostra; e o número relativamente pequeno de informantes (de modo que cada indivíduo tem um peso grande no resultado geral).<sup>21</sup> A análise de uma amostra maior de migrantes, futuramente, poderá indicar se de fato há diferenças significativas entre migrantes paraibanos e nativos do sudeste; por ora, para essas análises preliminares, tal resultado alerta para a necessidade de que se considerem os falantes individuais na interpretação de resultados.

### **2.4.3 Análises por comunidades**

Esta subseção apresenta os resultados de modelos multivariados, em que se comparam os padrões de realização de vogais médias pretônicas dentro de cada comunidade, de acordo com as variáveis independentes descritas na seção 2.4.1: Contexto Fônico Precedente, Contexto Fônico Seguinte, F1 da Vogal da Sílabla Seguinte, Vogal Tônica, Distância da Vogal Tônica, Estrutura da Sílabla Pretônica, Sexo/Gênero, Idade de Chegada e Anos na Capital (as duas últimas, apenas nas amostras de paraibanos).

#### **Rio de Janeiro: PEUL vs. PBRJ**

Na Amostra PEUL, a análise multivariada sobre a vogal pretônica /e/ em

---

<sup>21</sup>De fato, em análise similar em que não se incluiu a variável Informante como efeito aleatório, a variável Amostra se apresenta como significativamente correlacionada com as medidas de F1. Tal resultado, na verdade, só reforça a importância de se incluir o Informante como efeito aleatório em todas as análises, a fim de que não se chegue a generalizações equivocadas, mesmo que estejam de acordo com as expectativas iniciais.

um modelo de efeitos mistos não selecionou qualquer variável independente como significativamente correlacionada com a altura da vogal, o que indica a preponderância do papel dos indivíduos e dos itens lexicais para explicar a variabilidade. Na Amostra PBRJ, há correlação somente com a variável F1 da Vogal da Sílabla Seguinte.<sup>22</sup>

**Tabela 7:** Estimativas (em Hz) de valores de F1 para vogais /e/ pretônicas de acordo com F1 da sílabla seguinte, na amostra PBRJ (N = 250)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas</b>					
( <i>Intercept</i> )	472,68	62,502	7,563	<0,001	***
F1.SIL.SEG	0,204	0,049	4,183	<0,001	***
<b>Vogais normalizadas</b>					
( <i>Intercept</i> )	349,079	13,148	26,549	<0,001	***
F1.SIL.SEG	0,044	0,012	3,83	<0,001	***

Na Tabela 7, o valor de *intercept* diz respeito ao valor de referência da variável independente; como se trata de uma variável contínua, tal estimativa avalia quanto seria o valor da variável dependente (medida de F1 da vogal pretônica /e/) quando a variável independente tem valor zero; a estimativa 0,204 Hz (para vogais não normalizadas) e 0,044 (para vogais normalizadas) para F1 da Vogal da Sílabla Seguinte indica que, a cada unidade dessa variável, estima-se que o valor da variável dependente aumentará nessa proporção. Por exemplo, se o valor de F1 da Vogal da Sílabla Seguinte for de 600 Hz, estima-se que o F1 da vogal pretônica /e/ seja de  $472,68 + (600 \times 0,204) = 595,08$

<sup>22</sup>Os resultados da Tabela 7 foram extraídos de modelos com a seguinte configuração no R: `lmer(F1 ~ SEXO + CONT.PREC + CONT.SEG + F1.SIL.SEG + VOGAL.TONICA + DIST.TONICA + ESTR.SIL.PRET + (1|INFORMANTE) + (1|Word), data=PBRJ.e)`. A única diferença para a análise de vogais normalizadas é a inclusão de `F1.norm` como variável dependente.

Hz. Em outras palavras, quanto mais baixa a realização da vogal da sílaba seguinte, mais baixa também tende a ser a realização da vogal pretônica. Tal resultado está de acordo com análises prévias sobre vogais pretônicas, que apontam para o favorecimento de realizações com o traço [-alto] quando a sílaba seguinte contém uma vogal com esse traço. A presente análise, no entanto, mostra que tal relação é sistemática e gradual: não se trata apenas de um processo fonológico abstrato (a saber, a harmonia vocálica), mas de um fenômeno fonético que atua a cada item lexical produzido pelos falantes: quanto mais baixa for a realização da vogal da sílaba seguinte, seja ela com F1 a 500 Hz, 600 Hz ou 700 Hz – mesmo que para um mesmo item lexical produzido diversas vezes –, mais baixa tenderá a ser a realização da vogal pretônica. É interessante notar que, para a vogal /e/ no Rio de Janeiro, tal fenômeno se aplica à fala dos paraibanos, mas não para os cariocas, o que caracteriza a diferença dialetal.

Por outro lado, quando se analisa a realização da vogal pretônica /o/, uma correlação significativa com F1 da Vogal da Sílaba Seguinte (e somente com essa variável) se dá na fala de cariocas (Tabela 8) e de paraibanos (Tabela 9). Para cariocas, a estimativa de aumento no valor de F1 da vogal pretônica é de 0,109 Hz para vogais não normalizadas e de 0,046 Hz para vogais normalizadas, a cada unidade de F1 da Vogal da Sílaba Tônica. Na fala de paraibanos, a estimativa de abaixamento da vogal pretônica é comparativamente bem maior: 0,367 Hz para vogais não normalizadas e 0,077 Hz para vogais normalizadas.

As correlações verificadas com F1 da Vogal da Sílaba Seguinte, exceto para a vogal pretônica /e/ na Amostra PEUL, reforçam a conclusão de



**Tabela 8:** Estimativas (em Hz) de valores de F1 para vogais /o/ pretônicas de acordo com F1 da sílaba seguinte, na amostra PEUL (N = 299)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas</b>					
( <i>Intercept</i> )	490,432	57,478	8,533	<0,001	***
F1.SIL.SEG	0,109	0,049	2,234	0,026	*
<b>Vogais normalizadas</b>					
( <i>Intercept</i> )	354,434	17,664	20,065	<0,001	***
F1.SIL.SEG	0,046	0,019	2,341	0,02	*

estudos prévios que apontaram essa variável como a mais importantes para a realização [-alta] de vogais pretônicas. Verifica-se que tal fenômeno também ocorre na fala de nativos do Rio de Janeiro (*pace* Yacovenco 1993), embora restrita à vogal /o/ e em um grau mais fraco do que na fala de migrantes paraibanos.

Contudo, chama a atenção a ausência de correlações significativas com as demais variáveis linguísticas e com as variáveis sociais, diferentemente do que verificaram Yacovenco (1993) e Marques (2006). A se valer pelo número de informantes das amostras, talvez não surpreenda a falta de correlação

**Tabela 9:** Estimativas (em Hz) de valores de F1 para vogais /o/ pretônicas de acordo com F1 da sílaba seguinte, na amostra PBRJ (N = 250)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas</b>					
( <i>Intercept</i> )	317,266	91,712	3,459	0,001	**
F1.SIL.SEG	0,367	0,059	6,267	<0,001	***
<b>Vogais normalizadas</b>					
( <i>Intercept</i> )	318,598	19,822	16,073	<0,001	***
F1.SIL.SEG	0,077	0,014	5,425	<0,001	***

com a variável Sexo/Gênero (nas Amostras PEUL e PBRJ), tampouco com Idade de Chegada e Anos na Capital para os paraibanos. A não correlação com variáveis linguísticas também pode se dever ao número relativamente reduzido de dados de cada amostra (entre 250 e 300), ao subdividi-las entre as vogais /e/ e /o/. No entanto, em modelos similares apenas com efeitos fixos (ou seja, com exclusão das variáveis Informante e Item Lexical), um número maior de variáveis sociais e linguísticas é selecionado, o que indica que os resultados das Tabelas 7 a 9 se devem ao controle das variáveis aleatórias. Desse modo, é possível que resultados e conclusões de estudos prévios tenham sido enviesados pela contribuição de poucos falantes ou itens lexicais, aos quais, até o momento, pouco se atentou.

### São Paulo: SP2010 *vs.* PBSP

Na Amostra SP2010, a testagem do mesmo modelo estatístico para a vogal pretônica /e/ revela um resultado distinto (ver Tabela 10): em comparação com os padrões de cariocas e de paraibanos no Rio de Janeiro, neste caso, correlaciona-se significativamente com a altura da vogal pretônica o Sexo/Gênero dos falantes.

**Tabela 10:** Estimativas (em Hz) de valores de F1 para vogais /e/ pretônicas de acordo com o sexo/gênero do falante, na amostra SP2010 (N = 350)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas</b>					
( <i>Intercept</i> )	549,229	55,243	9,942	<0,001	***
SEXOMasc	-114,44	19,828	-5,772	0,001	***
<b>Vogais normalizadas</b>					
( <i>Intercept</i> )	375,004	11,446	32,763	<0,001	***
SEXOMasc	-7,595	3,629	-2,093	0,037	*

Na Tabela 10, o valor de referência é o sexo feminino; em relação às mulheres, os homens paulistanos tendem a realizações mais altas da vogal pretônica /e/, com valores de F1 mais baixos. Tal resultado se verifica tanto com vogais não normalizadas (-114,4 Hz) quanto com vogais normalizadas (-7,6 Hz), o que indica que a variação inter-falantes devido a diferenças fisiológicas do aparelho acústico não pode dar conta dessas estimativas. Esse resultado surpreende porque se poderia esperar que, em caso de correlação com Sexo/Gênero, seriam os homens que favoreciam realizações [-alto] – à semelhança dos resultados de Yacovenco (1993) –, não as mulheres. Uma explicação para esse fato permanece em suspenso.

**Tabela 11:** Estimativas (em Hz) de valores de F1 para vogais /e/ pretônicas de acordo com o contexto fônico precedente, na amostra PBSP (N = 348)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas</b>					
( <i>Intercept</i> )	635,805	65,875	9,652	<0,001	***
CONT.PREClabial	15,317	33,562	0,456	0,649	
CONT.PRECvelar	59,486	61,173	0,972	0,333	
CONT.PRECPalatal	-17,715	46,054	-0,385	0,701	
CONT.PRECSibilante	37,55	35,627	1,054	0,294	
CONT.PRECVibrante	124,721	37,631	3,314	0,001	**
<b>Vogais normalizadas</b>					
( <i>Intercept</i> )	371,531	14,847	25,023	<0,001	***
CONT.PREClabial	4,54	7,627	0,595	0,553	
CONT.PRECvelar	11,419	14,041	0,813	0,418	
CONT.PRECPalatal	-7,828	10,357	-0,756	0,452	
CONT.PRECSibilante	5,959	8,127	0,733	0,466	
CONT.PRECVibrante	29,586	8,577	3,449	<0,001	***

Para os paraibanos em São Paulo, por sua vez, a realização da vogal pretônica /e/ se correlaciona significativamente com o Contexto Fônico Precedente, conforme apresentado na Tabela 11. Em relação ao valor de referência

**Tabela 12:** Estimativas (em Hz) de valores de F1 para vogais /o/ pretônicas de acordo com o contexto fônico precedente e F1 da sílaba seguinte, na amostra SP2010 (N = 350)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas</b>					
( <i>Intercept</i> )	437,427	58,287	7,505	<0,001	***
CONT.PREClabial	24,12	28,509	0,846	0,401	
CONT.PREClabial	17,725	32,41	0,547	0,586	
CONT.PREClabial	22,056	33,722	0,654	0,515	
CONT.PRECsibilante	167,216	45,88	3,645	<0,001	***
CONT.PREClabial	7,857	88,723	0,089	0,93	
F1.SIL.SEG	0,214	0,04	5,283	<0,001	***
<b>Vogais normalizadas</b>					
( <i>Intercept</i> )	356,708	11,818	30,184	<0,001	***
CONT.PREClabial	2,805	6,025	0,466	0,643	
CONT.PREClabial	3,922	6,851	0,572	0,569	
CONT.PREClabial	1,801	7,151	0,252	0,802	
CONT.PRECsibilante	30,108	9,824	3,065	0,003	**
CONT.PREClabial	0,85	19,602	0,043	0,965	
F1.SIL.SEG	0,04	0,009	4,333	<0,001	***

(consoantes dental-alveolares [t, d, n, l, r]), há diferenças significativas para as vibrantes posteriores (como em *relógio*), que favorecem a realização [-alta] da vogal pretônica (em média, +124,7 Hz para vogais não normalizadas e 29,6 Hz para as vogais normalizadas). Esse resultado concorda com aquele obtido por Pereira (2010), que verificou favorecimento do abaixamento justamente quando a vogal era precedida por esse segmento.

A mesma variável independente se mostra relevante para a realização da vogal pretônica /o/ na fala de paulistanos (Tabela 12). Neste caso, a diferença significativa ocorre entre as consoantes dental-alveolares e as consoantes sibilantes [s, z] (como em *sofá*), que relativamente favorecem a realização [-alto] em 167,2 Hz para vogais não normalizadas e em 30,1 Hz para

vogais normalizadas. Correlação semelhante foi observada em Pereira (2010) para seus informantes paraibanos. Além de Contexto Fônico Precedente, a medida de F1 da Vogal da Sílabla Seguinte se mostra correlacionada com a altura da vogal apenas para a realização da pretônica /o/, à semelhança do que se verificou para os cariocas. Há uma tendência à realização [-alta] em 0,214 Hz para vogais não normalizadas e 0,04 Hz para vogais normalizadas, a cada unidade da variável independente.

O português paulistano, ainda que normalmente não seja apontado como uma variedade em que ocorre abaixamento de vogais médias pretônicas, apresenta sim tendências semelhantes ao padrão de outras comunidades, com aplicação da regra nos mesmos contextos (para a vogal /o/, quando precedida de sibilante e seguida de vogal da sílabla seguinte com F1 alto); a diferença parece ser de grau, a ponto de o fenômeno passar despercebido pela consciência metalinguística dos falantes.

**Tabela 13:** Estimativas (em Hz) de valores de F1 para vogais /o/ pretônicas de acordo com F1 da sílabla seguinte, na amostra PBSP (N = 302)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas</b>					
( <i>Intercept</i> )	499,538	65,586	7,617	<0,001	***
F1.SIL.SEG	0,270	0,061	4,445	<0,001	***
<b>Vogais normalizadas</b>					
( <i>Intercept</i> )	353,08	18,482	19,104	<0,001	***
F1.SIL.SEG	0,057	0,018	3,172	0,002	**

Por fim, a Tabela 13 mostra que, para os paraibanos em São Paulo, apenas a medida de F1 da Vogal da Sílabla Seguinte afeta significativamente a altura da vogal pretônica /o/. Assim como na comparação entre cariocas e

paraibanos no Rio de Janeiro, tal tendência parece ser mais marcada para os migrantes do que para os nativos paulistanos (Cf. 0,270 Hz vs. 0,214 Hz; 0,057 Hz vs. 0,04 Hz). É possível, no entanto, que tais diferenças não sejam significativas.

	PEUL		PBRJ		SP2010		PBSP	
	/e/	/o/	/e/	/o/	/e/	/o/	/e/	/o/
Sexo/Gênero	-	-	-	-	✓	-	-	-
Idade de Chegada	NA	NA	-	-	NA	NA	-	-
Anos na Capital	NA	NA	-	-	NA	NA	-	-
Contexto Precedente	-	-	-	-	-	✓	✓	-
Contexto Seguinte	-	-	-	-	-	-	-	-
F1 da Vogal da Sílabla Seguinte	-	✓	✓	✓	-	✓	-	✓
Vogal Tônica	-	-	-	-	-	-	-	-
Distância da Tônica	-	-	-	-	-	-	-	-
Estrutura da Sílabla Pretônica	-	-	-	-	-	-	-	-

**Quadro 4:** Quadro resumo dos resultados de análises multivariadas sobre medidas de F1 de vogais médias pretônicas nas quatro amostras

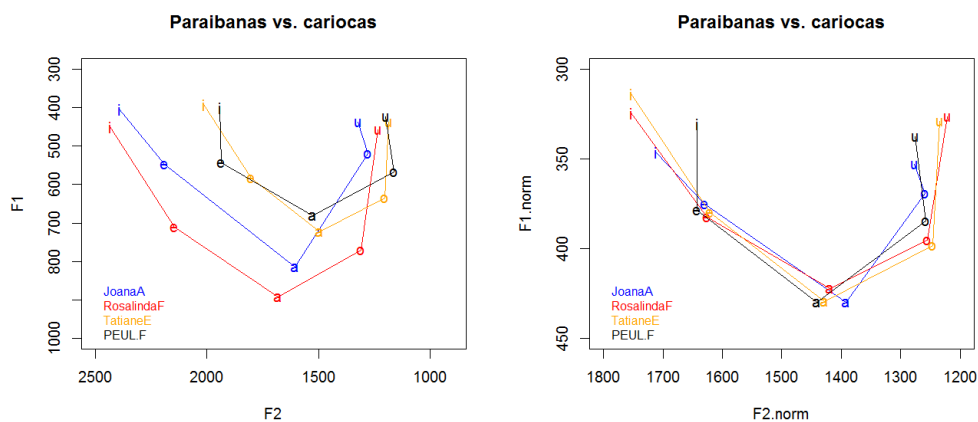
O Quadro 4 sintetiza os resultados das análises multivariadas sobre as vogais pretônicas /e/ e /o/ em cada uma das quatro amostras. Nela, percebe-se claramente que a variável F1 da Vogal da Sílabla Seguinte é a que se apresenta como mais importante para a altura das vogais médias pretônicas, tendo-se verificado correlações positivas em todos os casos em que revelou significância: quanto mais baixa a vogal da sílabla seguinte, mais baixa também tende a ser a vogal pretônica. O fenômeno ocorre para as vogais anterior e posterior apenas entre paraibanos no Rio de Janeiro e, nas demais comunidades, apenas para a vogal posterior. O Contexto Fônico Precedente se mostrou relevante para a realização de /o/ entre paulistanos e de /e/ entre paraibanos em São Paulo, em direções idênticas àquelas verificadas por

Pereira (2010). A variável social Sexo/Gênero, por fim, foi selecionada apenas na Amostra SP2010, com correlação em direção oposta à esperada (maior tendência a vogais [-alto] na fala de mulheres).

Cabe ressaltar que, em todos os modelos testados, foram selecionadas as mesmas variáveis independentes tanto na análise de vogais não normalizadas quanto de vogais normalizadas, o que dá maior sustentação aos resultados. Ademais, chama-se a atenção novamente para o pequeno número de correlações significativas, algo que parece ser decorrência mais do emprego de modelos de efeitos mistos, com variáveis aleatórias, do que do número de dados e de informantes (ainda que deva ser expandido). Diferentemente de estudos prévios, não se verificou correlação com, por exemplo, o Contexto Fônico Seguinte ou a Distância da Sílabas Tônicas, tampouco com o Sexo/Gênero dos falantes de modo geral. De fato, o comportamento individual de cada falante parece ter mais importância do que amplas categorias sociais. Contudo, como se trata de uma pequena amostra de informantes, volta-se a seguir para a análise da fala dos indivíduos, com vistas a levantar hipóteses para a expansão do *corpus*.

#### **2.4.4 Indivíduos**

As Figuras 6 a 9 abaixo apresentam o quadro vocálico de cada migrante paraibano, separados por Sexo/Gênero, em comparação com o padrão geral dos falantes nativos de cada comunidade. Igualmente à Figura 5 acima, apresenta-se o quadro com vogais não normalizadas à esquerda e com vogais normalizadas à direita. Nelas, a linha preta representa a fala de nativos do mesmo sexo, a fim de avaliar quais desses falantes mais se aproximam dos



**Figura 6:** Médias de F1/F2 das vogais pretônicas para paraibanas residentes no Rio de Janeiro, em comparação com cariocas (esq.: vogais não normalizadas; dir.: vogais normalizadas)

padrões da comunidade anfitriã.

Examine-se primeiramente a fala de mulheres paraibanas que residem no Rio de Janeiro (Figura 6). Há ampla variação entre elas quando se observam as vogais não normalizadas, mas seus quadros vocálicos se aproximam após a normalização. No quadro à direita, percebe-se que a realização da vogal pretônica /e/ é bastante semelhante entre as três informantes (JoanaA, RosalindaF e TatianeE), e entre elas e as mulheres cariocas. Por outro lado, elas parecem se diferenciar quanto à altura da vogal pretônica /o/.

Tais tendências foram analisadas em modelos de efeitos mistos, que incluíram o Item Lexical como efeito aleatório, e a variável Informante como efeito fixo, a fim de se obterem estimativas para cada uma dessas falantes. Nesse modelo, as três mulheres cariocas compõem o valor de referência.

Na Tabela 14, referente à vogal média [+anterior], observa-se que, embora a altura média dessas pretônicas não se diferencie significativamente quando se consideram as vogais normalizadas, tais diferenças emergem ao se analisarem



**Tabela 14:** Estimativas (em Hz) de valores de F1 para vogais /e/ pretônicas entre paraibanas no Rio de Janeiro, em comparação com cariocas (N = 300)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas<sup>a</sup></b>					
( <i>Intercept</i> )	540,128	8,543	63,224	<0,001	***
TatianeE	38,171	16,985	2,247	0,025	*
JoanaA	4,880	16,551	0,295	0,768	
RosalindaF	161,59	16,503	9,792	<0,001	***
<b>Vogais normalizadas<sup>b</sup></b>					
( <i>Intercept</i> )	378,108	2,332	162,105	<0,001	***
TatianeE	1,327	4,664	0,285	0,776	
JoanaA	-3,553	4,558	-0,78	0,436	
RosalindaF	2,205	4,532	0,486	0,627	

<sup>a</sup>Modelo: `lmer(F1~INFORMANTE+(1|Word), data=RJ.e[RJ.e$SEXO=="Fem"], )`

<sup>b</sup>Modelo: `lmer(F1.norm~INFORMANTE+(1|Word), data=RJ.e[RJ.e$SEXO=="Fem"], )`

as medidas de F1 para as vogais não normalizadas (presumivelmente, pela análise de Informante como efeito fixo, não como efeito aleatório). Das três paraibanas, JoanaA não apresenta diferenças significativas de realização de /e/ pretônico em comparação com as cariocas. De fato, no quadro superior esquerdo da Figura 6, é a sua vogal /e/ que mais se aproxima da altura das vogais das nativas. Por outro lado, as vogais pretônicas [+anterior] de TatianeE e de RosalindaF diferenciam-se pela realização mais baixa, que se reflete nas estimativas mais altas do modelo estatístico (+38,2 Hz e +161,6 Hz).

Quanto à vogal pretônica /o/ (Tabela 15), as três paraibanas têm estimativas significativamente distintas daquela para mulheres cariocas em vogais não normalizadas, e apenas JoanaA apresenta diferenças quando se consideram as vogais normalizadas. No entanto, enquanto TatianeE e RosalindaF apresentam tendência a realizar a vogal pretônica posterior com F1 mais alto

**Tabela 15:** Estimativas (em Hz) de valores de F1 para vogais /o/ pretônicas entre paraibanas no Rio de Janeiro, em comparação com cariocas (N = 299)

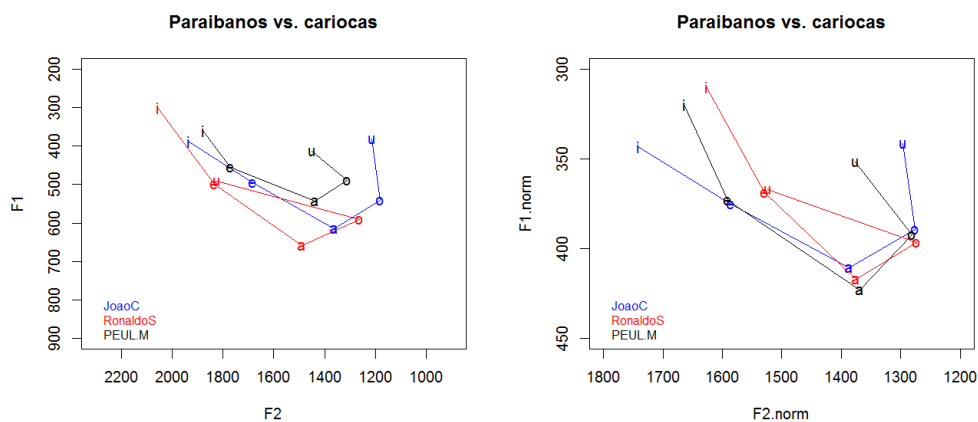
	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas<sup>a</sup></b>					
( <i>Intercept</i> )	573,8	12,171	47,143	<0,001	***
TatianeE	52,74	23,294	2,264	0,024	*
JoanaA	-60,631	22,273	-2,722	0,007	**
RosalindaF	186,462	22,233	8,387	<0,001	***
<b>Vogais normalizadas<sup>b</sup></b>					
( <i>Intercept</i> )	385,87	2,933	131,545	<0,001	***
TatianeE	10,429	5,768	1,808	0,072	
JoanaA	-17,729	5,526	-3,208	0,002	**
RosalindaF	7,007	5,504	1,273	0,204	

<sup>a</sup>Modelo: `lmer(F1~INFORMANTE+(1|Word), data=RJ.o[RJ.o$SEXO=="Fem"], )`

<sup>b</sup>Modelo: `lmer(F1.norm~INFORMANTE+(1|Word), data=RJ.o[RJ.o$SEXO=="Fem"], )`

(+52,7 Hz e +186,5 Hz), conforme o esperado para migrantes nordestinos, JoanaA revela, ao contrário, uma tendência a realizar o /o/ pretônico com F1 mais baixo (-60,6 Hz), ou seja, vogais pretônicas mais altas do que as das nativas. Poderia se afirmar que, em relação ao “alvo” da posição da vogal posterior, JoanaA não só o atinge, como o ultrapassa.

Impressionisticamente (para esta pesquisadora), JoanaA soa pouco nordestina – exceto em certos momentos da gravação, em que ressalta sua origem –, e apresenta outros traços tipicamente cariocas, como o /s/ em coda “chiado” [ʃ] e o *ingliding* de vogais tônicas. A maior acomodação de JoanaA ao padrão vocálico carioca pode estar relacionada com a idade à época de sua mudança para o Rio de Janeiro e com o tempo de residência na capital fluminense: a informante migrou no início da adolescência, aos 12 anos, sendo trazida para trabalhar em uma casa de família como empregada doméstica, e vive no Rio de Janeiro há 29 anos (ver Anexo B, p. 84). TatianeE e RosalindaF, por sua



**Figura 7:** Médias de F1/F2 das vogais pretônicas para paraibanos residentes no Rio de Janeiro, em comparação com cariocas (esq.: vogais não normalizadas; dir.: vogais normalizadas)

vez, migraram com 19 e 25 anos e vivem no Rio de Janeiro há 9 e 18 anos respectivamente. Embora as três tenham intenso contato com outros paraibanos (JoanaA é casada com o informante JoaoC desta amostra, Rosalinda também é casada com um paraibano, e TatianeE namora um conterrâneo), JoanaA também parece ser aquela mais integrada na comunidade e aquela que mais ascendeu socialmente desde que chegou ao Rio de Janeiro. Há cerca de 10 anos deixou de ser empregada doméstica e hoje obtém seu sustento como manicure; muitas de suas clientes são cariocas de classe média.

O quadro vocálico dos homens paraibanos no Rio de Janeiro se encontra na Figura 7.<sup>23</sup> Para eles, verificam-se igualmente diferenças significativas entre

<sup>23</sup>Nos quadros vocálicos dos homens (Figuras 7 e 9), por vezes se observa a localização da vogal [u] em posição central ou [+anterior], o que, evidentemente, é inesperado – ver, por exemplo, a localização da vogal /u/ para o informante Ronaldo S. O Prof. João Moraes, em comunicação pessoal, explicou que se trata de um problema do programa Praat, que muitas vezes não diferencia os formantes F1 e F2 (por ambos serem baixos nessa vogal), e acaba acusando um valor de F2 que, na verdade, é o de F3. Essas medições, que foram extraídas automaticamente com o *script* Vowel Analyzer, terão de ser corrigidas manualmente. As medidas de F1 e F2 para as vogais /e/ e /o/, no entanto, parecem ser confiáveis.

**Tabela 16:** Estimativas (em Hz) de valores de F1 para vogais /e/ pretônicas entre paraibanos no Rio de Janeiro, em comparação com cariocas (N = 250)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas<sup>a</sup></b>					
( <i>Intercept</i> )	460,737	6,106	75,458	<0,001	***
RonaldoS	44,821	9,838	4,556	<0,001	***
JoaoC	33,008	9,328	3,539	<0,001	***
<b>Vogais normalizadas<sup>b</sup></b>					
( <i>Intercept</i> )	374,585	2,306	162,422	<0,001	***
RonaldoS	-3,336	4,02	-0,83	0,407	
JoaoC	0,822	3,903	0,211	0,833	

<sup>a</sup>Modelo: `lmer(F1~INFORMANTE+(1|Word), data=RJ.e[RJ.e$SEXO=="Masc", ])`

<sup>b</sup>Modelo: `lmer(F1.norm~INFORMANTE+(1|Word), data=RJ.e[RJ.e$SEXO=="Masc", ])`

nativos e migrantes, tanto para a vogal /e/ (Tabela 16) quanto para a vogal /o/ (Tabela 17), mas apenas quando se consideram vogais não normalizadas. JoaoC, marido de JoanaA, apresenta menor tendência ao abaixamento de vogais pretônicas em relação a RonaldoS (Cf. 33 Hz vs. 44 Hz para /e/; 47,9 Hz vs. 100,9 Hz para /o/). Ambos migraram ao Rio de Janeiro em meados da adolescência, aos 17 e 16 anos respectivamente, mas lá vivem há períodos diferentes: JoaoC, hoje com 42 anos, vive no Rio de Janeiro há 25, enquanto RonaldoS, com 31, mudou-se há 15 anos. Neste caso, parece que a maior tendência à acomodação parece estar relacionada com o tempo de residência no sudeste, e não com a idade à época de migração.

No entanto, impressionisticamente, é curioso que RonaldoS soe (a esta pesquisadora) como mais “carioca” do que JoaoC. Traços como o “s” chiado e o *ingliding* de vogais tônicas também são frequentes em sua fala, de modo semelhante a JoanaA. A impressão de menor de grau de “nordestinidade”, desse modo, pode se dever a outros traços linguísticos que não as vogais

**Tabela 17:** Estimativas (em Hz) de valores de F1 para vogais /o/ pretônicas entre paraibanos no Rio de Janeiro, em comparação com cariocas (N = 250)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas<sup>a</sup></b>					
( <i>Intercept</i> )	489,368	6,783	72,149	<0,001	***
RonaldoS	100,939	12,156	8,303	<0,001	***
JoaoC	47,935	11,715	4,092	<0,001	***
<b>Vogais normalizadas<sup>b</sup></b>					
( <i>Intercept</i> )	392,926	2,502	157,068	<0,001	***
RonaldoS	2,969	4,569	0,65	0,516	
JoaoC	-4,972	4,417	-1,126	0,261	

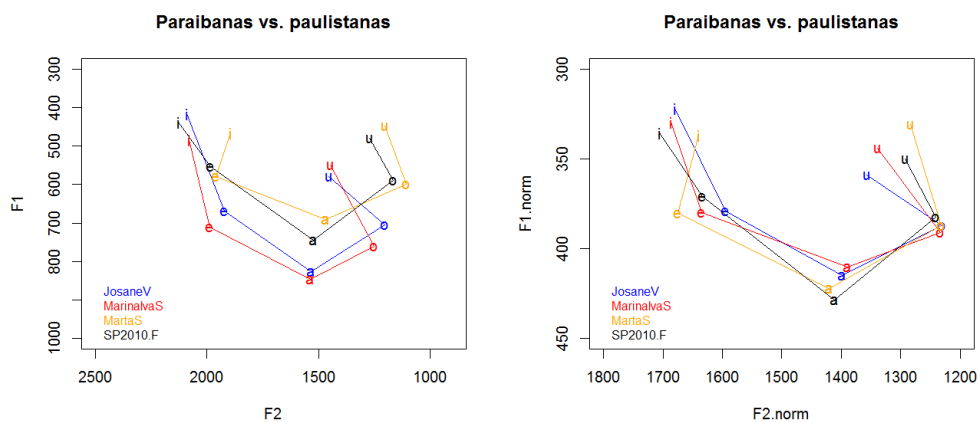
<sup>a</sup>Modelo: `lmer(F1~INFORMANTE+(1|Word), data=RJ.o[RJ.o$SEXO=="Masc", ])`

<sup>b</sup>Modelo: `lmer(F1.norm~INFORMANTE+(1|Word), data=RJ.o[RJ.o$SEXO=="Masc", ])`

pretônicas – o que reforça a necessidade de medidas objetivas de análise, como a medição de F1, em vez de procedimentos como a codificação de oitava, que poderiam ter enviesado os resultados.

Examinem-se em seguida as paraibanas na cidade de São Paulo (Figura 8). As três migrantes, JosaneV, MarinalvaS e MartaS, apresentam as vogais pretônicas /e/ e /o/ mais baixas do que as vogais das paulistanas, mas MartaS parece ser a que mais se aproxima do padrão das nativas.

Em modelo de regressão linear (novamente com Item Lexical como efeito aleatório e Informante como efeito fixo), tal tendência se confirma tanto para a vogal /e/ (Tabela 18) quanto para a vogal /o/ (Tabela 19): MartaS não apresenta diferenças significativas em relação ao valor de referência paulistano. JosaneV e MarinalvaS, em seu turno, exibem diferenças significativas para a vogal /e/ não apenas no modelo que incluiu as vogais não normalizadas, mas também naquele em que se analisaram as vogais normalizadas; em outras palavras, as diferenças na altura de F1 da vogal pretônica persistem mesmo



**Figura 8:** Médias de F1/F2 das vogais pretônicas para paraibanas residentes em São Paulo, em comparação com paulistanas (esq.: vogais não normalizadas; dir.: vogais normalizadas)

após descontar características individuais de trato vocálico. Para a vogal /o/, as diferenças persistem na análise de vogais normalizadas para MarinalvaS.

O que se pode aventar como motivo para os diferentes padrões de acomodação entre as paraibanas em São Paulo? Diferentemente da aparente motivação para o padrão de JoanaA, que pode ter se acomodado devido à migração no início da adolescência e aos muitos anos de residência no Rio de Janeiro, e para o padrão de JoaoC, que também vive há mais de 20 anos na nova comunidade, MartaS nem migrou cedo (mudou-se aos 18 anos) nem vive há em São Paulo por período comparável (há 14 anos). JosaneV e MarinalvaS, em contraste, vivem na capital paulista há mais tempo (16 e 23 anos), tendo se mudado igualmente ao fim da adolescência (aos 21 e 18 anos, respectivamente). Com efeito, MarinalvaS, mesmo sendo aquela que há mais tempo vive em São Paulo, é quem menos parece ter se acomodado por realizar as vogais pretônicas com medidas bastante elevadas de F1. As redes sociais das três são bastante semelhantes, com muitos amigos e parentes da

**Tabela 18:** Estimativas (em Hz) de valores de F1 para vogais /e/ pretônicas entre paraibanos em São Paulo, em comparação com paulistanos (N = 301)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas<sup>a</sup></b>					
( <i>Intercept</i> )	548,246	8,728	62,812	<0,001	***
MartaS	23,703	15,546	1,525	0,128	
JosaneV	116,566	15,841	7,359	<0,001	***
MarinalvaS	161,899	15,241	10,622	<0,001	***
<b>Vogais normalizadas<sup>b</sup></b>					
( <i>Intercept</i> )	369,153	2,386	154,684	<0,001	***
MartaS	8,058	4,215	1,912	0,057	
JosaneV	8,574	4,305	1,992	0,047	*
MarinalvaS	10,742	4,127	2,603	0,01	**

<sup>a</sup>Modelo: `lmer(F1~INFORMANTE+(1|Word), data=SP.e[SP.e$SEXO=="Fem", ])`

<sup>b</sup>Modelo: `lmer(F1.norm~INFORMANTE+(1|Word), data=SP.e[SP.e$SEXO=="Fem", ])`

Paraíba ou do Nordeste. As hipóteses anteriormente levantadas parecem não se aplicar a essas falantes.

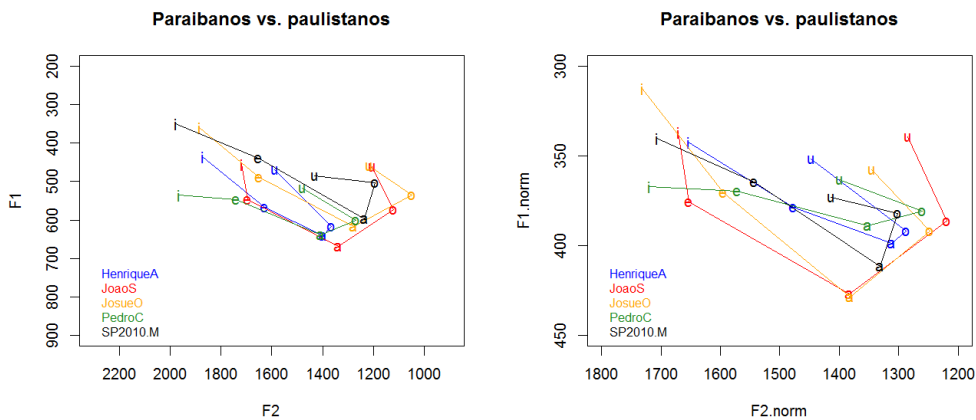
Dentre os paraibanos em São Paulo (ver Figura 9), todos exibem diferenças significativas em relação aos nativos tanto para a pronúncia da pretônica anterior (Tabela 20) quanto da pretônica posterior (Tabela 21), e ambas no sentido de realizações [-alta]. Mesmo na análise de vogais normalizadas, JoaoS e HenriqueA mantêm diferença significativa na altura da vogal pretônica /e/, e JosueO apresenta padrão divergente ao dos paulistanos para a vogal /o/. Relativamente aos demais, JosueO parece ser aquele que mais se acomodou ao padrão vocálico paulistano em valores absolutos (valores menores de estimativa em vogais não normalizadas). Este é o falante que mais parece ter ascendido socialmente em seus anos em São Paulo: trabalha atualmente em uma grande empresa como técnico de ar condicionado, e tem perspectivas de fazer faculdade (com incentivo financeiro de sua empresa) e se tornar

**Tabela 19:** Estimativas (em Hz) de valores de F1 para vogais /o/ pretônicas entre paraibanos em São Paulo, em comparação com paulistanos (N = 295)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas<sup>a</sup></b>					
( <i>Intercept</i> )	583,275	7,682	75,923	<0,001	***
MartaS	14,008	13,795	1,015	0,311	
JosaneV	110,836	14,419	7,687	<0,001	***
MarinalvaS	170,837	14,278	11,965	<0,001	***
<b>Vogais normalizadas<sup>b</sup></b>					
( <i>Intercept</i> )	380,884	2,187	174,176	<0,001	***
MartaS	5,72	3,793	1,508	0,133	
JosaneV	3,353	3,935	0,852	0,395	
MarinalvaS	7,822	3,924	1,993	0,047	*

<sup>a</sup>Modelo: `lmer(F1~INFORMANTE+(1|Word), data=SP.o[SP.o$SEXO=="Fem", ])`

<sup>b</sup>Modelo: `lmer(F1.norm~INFORMANTE+(1|Word), data=SP.o[SP.o$SEXO=="Fem", ])`



**Figura 9:** Médias de F1/F2 das vogais pretônicas para paraibanos residentes em São Paulo, em comparação com paulistanos (esq.: vogais não normalizadas; dir.: vogais normalizadas)



**Tabela 20:** Estimativas (em Hz) de valores de F1 para vogais /e/ pretônicas entre paraibanos em São Paulo, em comparação com paulistanos (N = 397)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas<sup>a</sup></b>					
( <i>Intercept</i> )	437,233	10,385	42,101	<0,001	***
JoaoS	109,926	21,131	5,202	<0,001	***
JosueO	47,273	20,903	2,262	0,024	*
HenriqueA	130,092	23,014	5,653	<0,001	***
PedroC	109,201	20,893	5,227	<0,001	***
<b>Vogais normalizadas<sup>b</sup></b>					
( <i>Intercept</i> )	364,89	2,32	157,264	<0,001	***
JoaoS	11,317	4,615	2,452	0,015	*
JosueO	5,234	4,562	1,147	0,252	
HenriqueA	14,309	5,085	2,814	0,005	**
PedroC	4,522	4,554	0,993	0,321	

<sup>a</sup>Modelo: `lmer(F1~INFORMANTE+(1|Word), data=SP.e[SP.e$SEXO=="Masc", ])`

<sup>b</sup>Modelo: `lmer(F1.norm~INFORMANTE+(1|Word), data=SP.e[SP.e$SEXO=="Masc", ])`

**Tabela 21:** Estimativas (em Hz) de valores de F1 para vogais /o/ pretônicas entre paraibanos em São Paulo, em comparação com paulistanos (N = 357)

	Estimativa	Erro padrão	Valor de <i>t</i>	Significância ( <i>p</i> )	
<b>Vogais não normalizadas<sup>a</sup></b>					
( <i>Intercept</i> )	506,884	11,289	44,901	<0,001	***
JoaoS	65,718	21,412	3,069	0,002	**
JosueO	47,896	21,026	2,278	0,023	*
HenriqueA	93,42	33,177	2,816	0,005	**
PedroC	96,127	20,685	4,647	<0,001	***
<b>Vogais normalizadas<sup>b</sup></b>					
( <i>Intercept</i> )	382,273	2,491	153,455	<0,001	***
JoaoS	0,514	5,21	0,099	0,921	
JosueO	13,986	5,075	2,756	0,006	**
HenriqueA	5,688	7,876	0,722	0,471	
PedroC	-2,223	4,999	-0,445	0,657	

<sup>a</sup>Modelo: `lmer(F1~INFORMANTE+(1|Word), data=SP.o[SP.o$SEXO=="Masc", ])`

<sup>b</sup>Modelo: `lmer(F1.norm~INFORMANTE+(1|Word), data=SP.o[SP.o$SEXO=="Masc", ])`

engenheiro. Trata-se também daquele que menos se identifica com o estado da Paraíba; ele diz não gostar de se lembrar de seu passado em Serra Branca, para onde volta para ficar apenas dois ou três dias em suas raras visitas a sua mãe, e diz se considerar, hoje, um paulistano. Por outro lado, quando se consideram as vogais normalizadas, é PedroC que não apresenta diferenças significativas em relação aos paulistanos. Assim como JoanaA, PedroC migrou cedo, com 13 anos, e vive na nova comunidade há 23. Este informante considera ter melhorado de vida consideravelmente desde que chegou há São Paulo, tendo já vivido em favelas mas hoje sendo dono de casa própria em bairro de classe média; é zelador da Faculdade de Letras da USP e tem, de acordo com seu discurso, a comunidade universitária como um exemplo que seus filhos paulistanos devem seguir. Questões de atitude quanto às comunidades parecem ter influência nos padrões dos paraibanos que residem em São Paulo.

#### **2.4.5 Conclusões parciais**

Todas as observações feitas acima a partir de exemplos individuais não permitem grandes generalizações, pelo pequeno número de falantes cuja fala foi até o momento analisada. Para cada uma das características apontadas para a relativa acomodação dos paraibanos – idade de migração, tempo de residência, ascensão social, atitudes –, há contraexemplos de outros falantes com características similares, mas que não se acomodaram no mesmo grau.

Entretanto, é possível traçar algumas considerações gerais. Percebe-se que há falantes que se aproximam bastante dos padrões de realização de vogais médias pretônicas dos nativos, como JoanaA (PBRJ) e MartaS

(PBSP), a ponto de não se observarem diferenças significativas em relação às vogais dos nativos das respectivas comunidades anfitriãs. Desse modo, os padrões linguísticos adquiridos na infância certamente podem ser modificados. À semelhança dos estudos reportados em Paiva & Duarte (2003), no Rio de Janeiro, e em Sankoff & Blondeau (2007), em Montréal, a respeito de indivíduos que foram gravados em duas épocas distintas, verifica-se a possibilidade de mudança, embora os fatores para tanto ainda não estejam claros.

Na presente amostra, os dois falantes que migraram no início da adolescência para as novas comunidades, JoanaA (com 12 anos) e PedroC (com 13 anos), estão entre os que mais se acomodaram aos padrões vocálicos do sudeste. Ainda que MartaS, que migrou aos 18, seja uma exceção a esta tendência, a idade à época da migração parece ter um papel importante.

O papel de tempo de residência, por outro lado, parece ser menos claro. Os informantes paraibanos dessas amostras residem desde há 9 até 25 anos no sudeste. Se se dividem os falantes entre aqueles que migraram há mais ou há menos de 20 anos, há aqueles que migraram há mais de 20 anos e se acomodaram mais (como PedroC, JoaoC e JoanaA); aqueles que migraram há mais de 20 anos e se acomodaram menos (como MarinalvaS e HenriqueA); os que migraram há menos de 20 anos e mais se acomodaram (como MartaS e JosueO); e os que migraram há menos de 20 anos e pouco se acomodaram (como JosaneV e RosalindaF). É a existência de falantes em todas as combinações possíveis de maior ou menor tempo de residência e maior ou menor acomodação que resulta em não correlação com essa variável independente.

Quanto às variáveis linguísticas, é interessante notar que a altura da

vogal da sílaba seguinte, segundo sua medição de F1, mostra-se como a mais revelante para a altura da vogal pretônica. Assim como em estudos prévios sobre a fala de nativos (Pereira, 2010) ou migrantes (Marques, 2006) paraibanos, esta variável se destaca. É especialmente digno de nota o fato de que essa variável também revela correlação significativa, no caso da vogal posterior /o/, na fala de cariocas e de paulistanos nativos, cujas variedades normalmente não se associam ao abaixamento de vogais médias pretônicas. Desse modo, este se apresenta como um condicionamento que parece ser válido para o PB de modo geral, e a diferença entre variedades ditas “do Norte” ou “do Sul” é mais uma questão quantitativa (de grau) do que qualitativa. Só foi possível observar esse resultado ao conceitualizar as vogais médias pretônicas como variáveis contínuas, que perpassam um amplo espectro de gradações em F1.

De modo semelhante, a correlação entre a altura da vogal pretônica e o Contexto Fônico Precedente na fala de paulistanos mostrou-se idêntica ao que se verifica na variedade paraibana (de acordo com Pereira 2010): em ambos os casos, são as consoantes [s, z] que favorecem a realização [-alta] da vogal pretônica, o que aponta, novamente, para condicionamentos gerais do PB.

Por fim, os modelos de efeitos mistos, com a análise sistemática de Informante e Item Lexical como efeitos aleatórios, alertam para a importância dessas variáveis em processos de variação e mudança linguística; diferentemente de estudos prévios, foram pouquíssimos os efeitos sociais e linguísticos que apresentaram qualquer correlação com a variável dependente. Ressalvado o fato de que a presente análise é restrita quanto ao número de informantes e

de ocorrências, ainda assim se assinala que certos condicionamentos teriam se mostrado significativos caso se utilizassem modelos de efeitos fixos (como no Varbrul). Estudos prévios, como o de Pereira (2010) e o de Marques (2006), já haviam discutido o papel de certos falantes e, sobretudo, de certos itens lexicais com comportamento idiossincrático ou categórico; na presente análise, contudo, foi possível incorporar tais fatos ao modelo estatístico, e não apenas avaliá-los de modo qualitativo.

## 2.5 Encaminhamentos futuros

As atividades aqui relatadas se referem a uma pesquisa em andamento, que terá continuidade pela pesquisadora na Unicamp. Trata-se, com efeito, apenas dos primeiros passos no estudo sistemático da fala de migrantes, tópico que ainda suscita mais perguntas do que respostas. Em especial, vislumbram-se os seguintes encaminhamentos:

- análise do papel de itens lexicais específicos: até o momento foi possível fazer uma análise mais sistemática de cada falante e de seus padrões de variação, conforme reportado na seção 2.4.4; análise semelhante deve ser feita sobre os itens lexicais, sobretudo aqueles que são mais frequentes.
- expansão da amostra: pelos presentes resultados, intenciona-se expandir a amostra de migrantes paraibanos controlando-se, além do sexo/gênero, a idade à época da migração e o tempo de residência na nova comunidade. Trata-se de duas variáveis que não são totalmente independentes entre si: é natural que aqueles que migraram mais cedo também sejam aqueles que vivem há mais tempo na comunidade anfitriã; o controle

sistemático dessas variáveis poderá desembaraçar o papel de cada uma delas.<sup>24</sup>

- coleta e observações mais sistemáticas sobre a classe e a rede social dos participantes: os comentários sobre possíveis motivos para a acomodação de falantes como JoanaA e JosueO levam em conta o fato de que são aqueles que parecem mais ter ascendido socialmente após a migração e mais ter se integrado à comunidade anfitriã; contudo, esses dados não foram sistematicamente anotados para cada informante. Para os novos informantes a serem gravados na expansão da amostra, provavelmente se incluirá algum tipo de questionário socioeconômico e perguntas específicas para mais bem avaliar o papel dessas características em processos de acomodação. A hipótese de Alves (1979) previa que falantes de nível socioeconômico mais alto teriam maior resistência à acomodação do que os migrantes mais pobres; aqui, no entanto, levanta-se outra hipótese: a de que os migrantes de baixo nível socioeconômico que ascenderam consideravelmente na nova comunidade tenderiam a maior grau de acomodação dialetal (ou seria seu maior grau de acomodação o facilitador de ascensão social?).
- aplicação dos mesmos métodos de análise a subamostra do ValPB: até o momento, os resultados das análises quantitativas foram comparados aos da pesquisa de Pereira (2010); entretanto, uma vez que os procedimentos de análise de vogais pretônicas foram estabelecidos (ver Anexo A, p. 82), sua aplicação a uma amostra de falantes paraibanos não migrantes

---

<sup>24</sup>Ou, alternativamente, também se pode analisar a variável tempo de residência como uma proporção da vida do falante, de modo semelhante a Bieler da Silva (2015).

permitirá uma comparação mais direta com os presentes resultados e estabelecerá um novo parâmetro para medir o grau de acomodação dos migrantes no sudeste.

Para além desses objetivos mais imediatos, a presente pesquisa deverá prosseguir durante vários anos porvir, tanto pela expansão das amostras para migrantes de outras localidades, quanto pela análise de novas variáveis dependentes (p.ex., palatalização de /ti, di/, pronúncia de /r/ em coda, duração de vogais etc.) Esta pesquisa realizada em nível de pós-doutoramento dá continuidade à análise iniciada no doutorado sobre a fala de moradores em São Paulo, e terá continuidade na exploração dos padrões sociolinguísticos na fala de outros membros das grandes metrópoles.

## 3 Atividades adicionais realizadas no período

### 3.1 Minicursos e palestras

- A variação em percepções sociolinguísticas. Palestra aos alunos do curso de graduação em Letras da Universidade Federal de São Paulo (Campus Guarulhos), organizada pelo InFoLinC (19/jun/2015). Carga horária: 2 horas.
- Sociolinguística Variacionista: novas tecnologias de análises de dados. Oficina ministrada durante o VI Seminário de Sociolinguística na Universidade Federal do Espírito Santo, organizado pelas professoras Maria Marta Pereira Scherre, Lilian Yacovenco e Leila Tesch (19/ago/2015). Carga horária: 2 horas.
- Introdução ao uso da plataforma R para a realização de análises estatísticas de dados linguísticos. Oficina ministrada na Universidade Federal do Rio de Janeiro, organizada pelas professoras Christina A. Gomes e Conceição Paiva (out-nov/2015). Carga horária: 18 horas.
- Introdução ao Rbrul. Oficina ministrada na Universidade Federal do Rio de Janeiro, organizada pelas professoras Silvia Figueiredo Brandão e Silvia Rodrigues Vieira (25–26/jan/2016). Carga Horária: 9 horas.
- Introdução ao R para análises estatísticas na Sociolinguística. Oficina ministrada na Universidade Federal da Paraíba, organizada pelo professor Dermeval da Hora (15–19/02/2016). Carga horária: 18 horas.



### 3.2 Participação em congressos

- OUSHIRO, L. Produção e percepção linguística sobre a pronúncia de (-r) em São Paulo. 2015. VI Seminário de Sociolinguística (SOLIN), Vitória, UFES, ago/2015.
- FOLTRAN, M. J.; NOBREGA, V. A.; OUSHIRO, L. Multiple determiners in Brazilian Portuguese indefinite noun phrases: intensification reading and agreement marker distribution. 2015. III Congresso Internacional de Estudos Linguísticos (CIEL), Brasília, UnB, ago/2015.
- OUSHIRO, L. Social meanings of (-r) in São Paulo: a computational approach for modeling the indexical field. 2015. New Ways of Analyzing Variation 44 (NWAV), Toronto-Canadá, out/2015.
- OUSHIRO, L.; MENDES, R. B. Diverging social perceptions: coda (-r) and variable number agreement in São Paulo Portuguese. 2015. New Ways of Analyzing Variation 44 (NWAV), Toronto-Canadá, out/2015.
- OUSHIRO, L. Modelagem do campo indexical da pronúncia variável do (-r) em São Paulo. 2015. Encontro Intermediário do GT de Sociolinguística, Porto Alegre, PUCRS, nov/2015.

### 3.3 Publicações

- OUSHIRO, L. Social and structural constraints in lectal cohesion. *Lingua* 172-3, p. 116-130, 2016. <http://www.sciencedirect.com/science/article/pii/S0024384115002181>

- MENDES, R. B.; OUSHIRO, L. Variable Number Agreement in Brazilian Portuguese: An Overview. *Language and Linguistics Compass*, v. 9, p. 358–368, 2015. <http://onlinelibrary.wiley.com/doi/10.1111/lnc3.12156/abstract>
- OUSHIRO, L. Dois pastel e um chopes: a concordância nominal e identidade(s) paulistana(s). *Revista de Estudos da Linguagem*, v. 23, p. 389–424, 2015. <http://periodicos.letras.ufmg.br/index.php/relin/article/view/6249>
- OUSHIRO, L. O que se diz e como se fala: relações entre o discurso metalinguístico e a variação linguística. *Signo y Señal - Revista del Instituto de Lingüística*, v. 28, p. 139–167, 2015. <http://revistas.filo.uba.ar/index.php/sys/article/view/304>
- OUSHIRO, L. Interação entre sexo/gênero e classe social no uso variável da concordância verbal. In: Freitag, Raquel Meister Ko; Severo, Cristine Gorski (orgs.). *Mulheres, Linguagem e Poder: estudos de gênero na sociolinguística brasileira*. 1ª ed. São Paulo: Blucher, 2015, p. 151–168. <http://pdf.blucher.com.br/openaccess/9788580391213/0006.pdf>
- OUSHIRO, L.; GUY, G. R. The effect of salience on co-variation in Brazilian Portuguese. *University of Pennsylvania Working Papers in Linguistics*, v. 21(2), p. 155–166, 2015. <http://repository.upenn.edu/pwpl/vol21/iss2/18/>

### **3.4 Participação em bancas de qualificação**

- Mendes, R. B.; Chagas de Souza, P.; Oushiro, L. Participação em banca de Larissa Grasiela Mendes Soriano. Percepções sociolinguísticas sobre o /-r/ em São Paulo. Exame de qualificação (Mestranda em Linguística), ago/2015 – Universidade de São Paulo.
- Chagas de Souza, P.; Sandalo, F.; Oushiro, L. Participação em banca de Júlia Sales Paez Fernandez. Por que seria o dinamarquês uma língua tão complicada? Processos de enfraquecimento consonantal na língua dinamarquesa. 2015. Exame de qualificação (Mestranda em Linguística), set/2015 – Universidade de São Paulo.

### **3.5 Reuniões com alunos de graduação e de pós-graduação**

Durante o segundo semestre de 2015, a pós-doutoranda deu atendimento aos alunos Thiago Laurentino de Oliveira (orientadora: Célia Regina Lopes), Christine Ministher (orientadora: Maria Cecilia Mollica), Isabella Coutinho Costa (orientadora: Kristine S. Stenzel), Tatiane Sudré e Guilherme Rodrigues (orientadora: Ana Paula Quadros Gomes), auxiliando-os em suas respectivas pesquisas com questões relativas a coleta de dados, elaboração de experimentos e análises estatísticas.

### **3.6 Emissão de pareceres**

- Revista Signótica. Parecer a artigo científico.
- Revista Todas as Letras. Parecer a artigo científico.

- Revista Linguística Rio. Parecer a artigo científico.
- I Simpósio Internacional de Actitudes ante Español, Portugués y Lenguas Relacionadas. Participação do Comitê Científico para emissão de pareceres sobre os resumos submetidos ao simpósio.
- New Ways of Analyzing Variation 44. Participação do Comitê Científico para emissão de pareceres sobre os resumos submetidos ao congresso.
- III Colóquio Brasileiro de Morfologia. Participação do Comitê Científico para emissão de pareceres sobre os resumos submetidos ao colóquio.
- Linguistic Symposium on Romance Languages 45. Parecer a artigo científico submetido para publicação de trabalhos selecionados.

## Referências

- ALVES, Maria Isolete Pacheco Menezes. *Atitudes linguísticas de nordestinos em São Paulo: abordagem prévia*. Dissertação de Mestrado. Campinas: IEL/Unicamp, 1979. 226f.
- BATTISTI, Elisa. *Elevação das vogais médias pretônicas em sílaba inicial de vocábulo na fala gaúcha*. Dissertação de Mestrado. Porto Alegre: UFRGS, 1993. 125f.
- BIELER DA SILVA, Mariane Esteves. *Entre duas metrópoles: (-r) em Itanhandu*. Dissertação de Mestrado. São Paulo: FFLCH-USP, 2015.
- BISOL, Leda. *Harmonização vocálica: uma regra variável*. Tese de Doutorado. Rio de Janeiro: UFRJ., 1981.
- BORTONI-RICARDO, Stella Maris. *The urbanization of rural dialect speakers*. Cambridge: Cambridge University Press, 1985.
- CALLOU, Dinah, LEITE, Yonne & COUTINHO, Lilian. “Elevação e abaixamento das vogais pretônicas no dialeto do Rio de Janeiro.” *Organon*, vol. 18, 71–78, 1991.
- CARDOSO, Daisy Barbara Borges. *Variação e mudança do imperativo no português brasileiro: gênero e identidade*. Tese de Doutorado. Brasília: Universidade de Brasília, 2009. 165f.
- CARMO, Maria Cristina do. *As vogais médias pretônicas na variedade do interior paulista*. Tese de Doutorado. São José do Rio Preto: UNESP, IBILCE, 2013.
- CELIA, Gianni Fontis. *As vogais médias pretônicas na fala culta de Nova Venécia*. Dissertação de Mestrado. Campinas: IEL/Unicamp, 2004.  
URL <http://www.bibliotecadigital.unicamp.br/document/?code=vtls000316849>

- CHACON, Karoline de Albuquerque. *Contato dialetal: análise do falar paulista em João Pessoa*. Dissertação de Mestrado. João Pessoa: Universidade Federal da Paraíba, 2012.
- CHAVES, Idalena Oliveira. *Panorama dos estudos das vogais pretônicas no Português do Brasil: meta-análise das pesquisas desenvolvidas de 1980 a 2012*. Tese de Doutorado. Belo Horizonte: Universidade Federal de Minas Gerais, 2013.
- CRUZ, Regina. “Vogais na Amazônia Paraense.” *Alfa*, vol. 56(3), 945–972, 2012.
- GOLDMAN, Jean-Philippe. “EasyAlign: an automatic phonetic alignment tool under Praat.” *In: Proceedings of InterSpeech*, 2011.  
URL <http://latlcul.unige.ch/phonetique/easyalign.php>
- HARRINGTON, Jonathan. “An acoustic analysis of ‘happy-tensing’ in the Queen’s Christmas broadcast.” *Journal of Phonetics*, vol. 34, 439–457, 2006.
- HARRINGTON, Jonathan, PALETHORPE, Sallyanne & WATSON, Catherine I. “Does the Queen speak the Queen’s English?” *Nature*, vol. 408, 927–928, 2000a.
- . “Monophthongal vowel changes in Received Pronunciation: an acoustic analysis of the Queen’s Christmas broadcasts.” *Journal of the International Phonetic Association*, vol. 30(1/2), 63–78, 2000b.
- . “Deepening or lessening the divide between diphthongs? An analysis of the Queen’s Annual Christmas broadcasts.” *In: A figure of speech: A Festschrift for John Laver*, pp. 227–261. Mahwah, NJ: Lawrence Erlbaum, 2005.
- HELLWIG, Birgit & GEERTS, Jeroen. “ELAN – Linguistic Annotator. Versão 4.4.0.”, 2013. Disponível em <http://www.mpi.nl/corpus/manuals/manual-elan.pdf>.

- HORA, Dermeval da & WETZELS, Leo. “Róticos: uma “po[h, r, ʝ]ta” entre paraibanos e paulistanos.” *Linguística*, vol. 24, 51–76, 2010.
- . “A variação linguística e as restrições estilísticas.” *Revista da ABRALIN*, vol. Eletrônico(Especial), 147–188, 2011.
- KENDALL, Tyler & THOMAS, Erik R. “Package ‘vowels’.” Disponível em <http://blogs.uoregon.edu/vowels/2012/11/08/vowels-r-1-2/>, 2015. Pacote para a plataforma R.
- LABOV, William. “The anatomy of style-shifting.” *In: ECKERT, P. & RICKFORD, J. R. (Eds.), Style and Sociolinguistic Variation*, pp. 85–108. Cambridge: Cambridge University Press, 2001*a*.
- . *Principles of linguistic change: social factors*. Oxford & Cambridge: Blackwell, 2001*b*.
- LEITE, Cândida M. Britto. *Atitudes linguísticas: a variante retroflexa em foco*. Dissertação de Mestrado. Campinas: Unicamp, 2004. 138f.
- LIMA, Izete de Souza & LUCENA, Rubens Marques de. “Influência de variáveis não linguísticas no processo de acomodação dialetal do /s/ em coda silábica por paraibanos em Recife.” *Letrônica*, vol. 6(1), 161–178, 2013.
- LOBANOV, Boris M. “Classification of Russian vowels spoken by different speakers.” *Journal of Acoustic Society of America*, vol. 49(2), 606–608, 1971.
- MARQUES, Sandra Maria Oliveira. *As vogais médias pretônicas em situação de contato dialetal*. Tese de Doutorado. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2006. 161f.
- MARTINS, Mariana de Souza. *A palatalização de oclusivas dentais em contato dialetal*. Dissertação de Mestrado. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2008. 146f.

- MENDES, Ronald Beline. “A pronúncia retroflexa do /-r/ na fala paulistana.”  
*In: HORA, D. & NEGRÃO, E. V. (Eds.), Estudos da Linguagem. Casamento entre temas e perspectivas*, pp. 282–299. João Pessoa: Ideia, 2011.
- MENDES, Ronald Beline & OUSHIRO, Livia. “O paulistano no mapa sociolinguístico brasileiro.” *Alfa*, vol. 56(3), 973–1001, 2012.  
 URL <http://seer.fclar.unesp.br/alfa/article/view/4942/4369>
- NASCENTES, Antenor. *O linguajar carioca*. Rio de Janeiro: Organização Simões, 1953 [1922].
- NYCZ, Jennifer R. *Second dialect acquisition: implications for theories of phonological representation*. Tese de Doutorado. Nova York: New York University, 2011. 254f.
- OLIVEIRA, Antônio Tadeu Ribeiro de, ERVATTI, Leila Regina & O’NEILL, Maria Monica Vieira Caetano. “O panorama dos deslocamentos populacionais no Brasil: PNADs e censos demográficos.” *In: OLIVEIRA, Luiz Antonio Pinto de & OLIVEIRA, Antônio Tadeu Ribeiro de (Eds.), Reflexões sobre os deslocamentos populacionais no Brasil*, vol. Estudos e análises: Informação demográfica e socioeconômica, pp. 28–48. Rio de Janeiro: IBGE/Ministério do Planejamento, Orçamento e Gestão, 2011.
- OUSHIRO, Livia. “Transcrição de entrevistas sociolinguísticas com o ELAN.”  
*In: FREITAG, Raquel Meister Ko (Ed.), Metodologia de coleta e manipulação de dados em Sociolinguística*. São Paulo: Blucher, 2014.  
 URL <http://blucheropenaccess.com.br/articles/download/290>
- . “silac: Silabificador, acentuador e transcritor fonológico do Português Brasileiro - v0.2.” Ms., 2015. (script do R).
- PAIVA, Maria da Conceição de & DUARTE, Maria Eugênia Lamoglia (Eds.). *Mudança linguística em tempo real*. Rio de Janeiro: Contra Capa, 2003.
- PEREIRA, Regina Celi Mendes. *Uma análise variacionista das vogais médias pretônicas*. João Pessoa: Editora da UFPB, 2010.



- R CORE TEAM. “R: A language and environment for statistical computing.” R Foundation for Statistical Computing, Vienna, Austria, 2015. Último acesso em 25 set./2014.  
URL <https://www.R-project.org/>
- RIEBOLD, John. “Vowel analyzer.” Ms., 2013. (script do Praat).  
URL <https://raw.githubusercontent.com/jmriebold/Praat-Tools/master/Vowel-Analyzer.praat>
- SANCIER, Michele L. & FOWLER, Carol A. “Gestural drift in a bilingual speaker of Brazilian Portuguese and English.” *Journal of Phonetics*, vol. 25, 421–436, 1997.
- SANKOFF, Gillian & BLONDEAU, Hélène. “Language change across the lifespan: /r/ in Montreal French.” *Language*, vol. 83(3), 560–588, 2007.
- SCHMIDT, Thomas & WORNER, Kai. “EXMARaLDA.” In: DURAND, Ulrike Gut Jacques & KRISTOFFERSEN, Gjert (Eds.), *Handbook on Corpus Phonology*, pp. 402–419. Oxford: Oxford University Press, 2014.
- SILVA, Myrian Barbosa da. *As pretônicas no falar baiano: a variedade culta de Salvador*. Tese de Doutorado. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 1989. 371f.
- SOARES, Viviane dos Ramos. *A negação no contato entre dialetos*. Dissertação de Mestrado. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2009. 114f.
- WEINREICH, Uriel, LABOV, William & HERZOG, Marvin I. *Fundamentos empíricos para uma teoria da mudança linguística*. São Paulo: Parábola, 2006 [1968]. Tradução de Marcos Bagno.
- YACOVENCO, L. C. *As vogais médias pretônicas no falar culto carioca*. Tese de Doutorado, Rio de Janeiro: UFRJ, 1993.
- ZANI, Juliana Camargo. *O alçamento das vogais médio-baixas no falar da cidade de São Paulo*. Dissertação de Mestrado. São Paulo: FFLCH-USP, 2009.

## Anexo A: Procedimentos Metodológicos

1. **Audacity:** Abrir arquivo de gravação no Audacity<sup>25</sup> e fazer duas mudanças: mudar para mono (se gravação estiver em stereo/2 canais) e downsample para 16000Hz (Faixas > ...). Salvar nova versão da gravação, que será usada no Praat, em pasta chamada Alinhamento, dentro de cada subpasta de corpus.
2. **ELAN:** Exportar trilha apenas da fala do informante (S1) para formato .TextGrid
3. **Praat:** Abrir arquivo exportado para formato .TextGrid e rodar passo 3 do *plugin* EasyAlign (Phone Segmentation); após rodá-lo (leva-se cerca de 30 minutos para uma gravação de 60 minutos), duplicar trilha “words” e nomear como “marc.pretonicas”
4. **ELAN:** Importar arquivo .TextGrid (como encoding UTF-16) e exportar todas as trilhas no formato texto delimitado por tabulador, UTF-8, tempo em milisegundos, adicionando “temp-” ao início do nome do arquivo
5. **Calc:** Abrir arquivo “temp-” exportado e salvá-lo no formato .csv
6. **R:** Abrir *script* silacpret.R; definir como diretório de trabalho (função `setwd()`) a pasta em que está o arquivo .csv e rodar o *script*; ele gerará um arquivo no formato .txt com adição de “segm-” ao nome original
7. **ELAN:** Abrir arquivo “segm-temp-...” no ELAN; exportar no formato .TextGrid para a pasta VowelAnalyzer-Praat, apagando “segm-temp-...” do nome para que arquivo fique com o mesmo nome que o arquivo de gravação
8. **Praat:** Rodar *script* do Vowel Analyzer; a trilha de vogais (Vowel Tier) é a “fones.pretonicas” (leva-se cerca de 30 minutos para extrair os dados de 6 arquivos).

---

<sup>25</sup>Disponível gratuitamente em <http://www.audacityteam.org/>. Último acesso em 19/04/2016.

9. **Bloco de Notas:** Abrir o arquivo gerado pelo Vowel Analyzer e salvar em formato UTF-8
10. **Calc:** Abrir arquivo salvo em formato UTF-8 e salvar em formato .csv
11. **R:** Abrir *script* pos-VowelAnalyzer.R (*script* para codificação automática de variáveis); definir o diretório de trabalho, nome do arquivo, nome dos informantes e rodá-lo; o *script* gerará uma planilha de dados
12. O arquivo de dados está pronto para ser analisado!

Tabela 22: Anexo B: Características sociais dos informantes

Amostra	Pseudônimo	Sexo	Idade	Escolaridade	Local de origem	Migração	Anos SE	Profissão	Bairro
PBRJ	RosalindaF	F	43	Fund. I	Capim de Momonguape	25	18	doméstica	Manguinhos
	JoanaA	F	41	Fund. I	Lagoa de Dentro	12	29	manicure	Manguinhos
	TatianeE	F	29	analfabeta	Sapé	19	9	ajudante em van	Manguinhos
	RonaldoS	M	31	Fund. II	Campina Grande	16	15	porteiro	Madureira
	JoãoC	M	42	Fund. I	Lagoa de Dentro	17	25	técnico de bateria	Manguinhos
	MartaS	F	32	Fund. I	Casserengue	18	14	doméstica	Patriarca
	JosaneV	F	37	Médio	Solânea	21	16	costureira	-
PBSP	MarinalvaS	F	39	Fund. I	Picuí	18	21	ajudante geral	-
	JosuéO	M	30	Médio	Serra Branca	15	15	técnico de AC	Tatuapé
	JoãoS	M	31	Fund. I	Campina Grande	14	17	porteiro	Campo Limpo
	HenriqueA	M	42	Médio	Solânea	17	25	motorista	-
	PedroC	M	36	Fund. I	Teixeiras	13	23	zelador	Jd. Bonfiglioli
	16-Car	M	48	Fund. II	Rio de Janeiro	NA	NA	pintor	Copacabana
	17-Sim	F	27	Fund. II	Rio de Janeiro	NA	NA	doméstica	Horto
PEUL	18-Luc	F	49	Fund. II	Rio de Janeiro	NA	NA	-	Vila da Penha
	19-Jor	M	37	Médio	Rio de Janeiro	NA	NA	-	Vila Isabel
	20-Rei	M	47	Fund. II	Rio de Janeiro	NA	NA	-	Marechal Hermes
	22-Ana	F	34	Médio	Rio de Janeiro	NA	NA	telefonista	Engenho Novo
	AliceM	F	30	Médio	São Paulo	NA	NA	gerente de vendas	Vila Formosa
	AnaS	F	32	Médio	São Paulo	NA	NA	dona de casa	Pirituba
	RenataC	F	35	Médio	São Paulo	NA	NA	gerente administrativo	Bela Vista
SP2010	RobsonF	M	31	Médio	São Paulo	NA	NA	vendedor	Vila Andrade
	NelsonF	M	37	Médio	São Paulo	NA	NA	porteiro	Cidade Ademar
	MaurícioB	M	36	Médio	São Paulo	NA	NA	pintor	Mooca
	LucianoT	M	39	Médio	São Paulo	NA	NA	promoter	Perdizes